AD NUMBER

AD004641

CLASSIFICATION CHANGES

TO: unclassified

FROM: secret

LIMITATION CHANGES

TO:

Approved for public release, distribution unlimited

FROM:

Controlling DoD Organization: Department of the Army, Attn: Public Affairs Office, Washington, DC 20310.

AUTHORITY

26 May 1964, Group 4, DoDD 5200.10, 26 July 1962; St-a per ESD, USAF ltr 27 May 80

AD-

4641

SECRET

NOISE-LIKE SIGNALS AND THEIR DETECTION BY CORRELATION

BENNETT L. BASORE

26 MAY 1952

TECHNICAL REPORT NO. 7

RESEARCH LABORATORY OF ELECTRONICS

and

LINCOLN LABORATORY

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

RESEARCH LABORATORY OF ELECTRONICS
AND
LINCOLN LABORATORY

NOISE-LIKE SIGNALS AND THEIR DETECTION BY CORRELATION

Bennett L. Basore

Technical Report No. 7

26 May 1952

CAMBRIDGE MASSACHUSETTS

NOISE-LIKE SIGNALS AND THEIR DETECTION BY CORRELATION*

ABSTRACT

Communication systems, in which a noise-like signal is used as an information carrier and cross-correlation at the receiver is used for detection, are investigated. The probability of error in the reception of signals by such systems (called NOMAC systems) is given as a function of input signal-to-noise ratio, input-to-output bandwidth ratio, and the number of possible signals. The effect of having a noisy version of the signal with which the input signal is cross-correlated is included, and the effect of using an arbitrary threshold value of the output as a criterion of detection is shown to result in a loss of available channel capacity, and correspondingly higher probability of error. Proposed practical systems employing NOMAC principles are described in some detail, along with the experimental system that has been constructed and tested. The experimental results are shown to agree with the theory.

*This report is identical with a thesis of the same title submitted in partial fulfillment of the requirements for the Degree of Doctor of Science in the Department of Electrical Engineering at the Massachusetts Institute of Technology, 26 May 1952.

CONTENTS

	ABSTRACT	ii
CHAPTER		
I.	THE BASIC COMMUNICATION SYSTEM	1
	A. Elements of Communication Theory	1
	B. The Mathematical Model	4
II.	CORRELATION DETECTION CRITERIA	9
III.	DESCRIPTION OF THE FUNDAMENTAL NOMAC SYSTEM	13
IV.	THE THEORETICAL STUDY OF PROBABILITY OF ERROR	17
	A. The Criterion of Maximum Correlation	17
	B. The Density Distribution of X_z	19
V.	THE THEORETICAL STUDY OF THRESHOLD DETECTION	27
	A. The Probability of Error for an Arbitrary Threshold	27
	B. The Optimum Threshold	31
VI.	NOISE IN THE AUXILIARY CHANNELS	35
VII.	THE EXPERIMENTAL STUDY OF PROBABILITY OF ERROR	39
	A. The Modified Theory for K = 1	39
	B. Description of the Equipment	41
	C. Discussion of Experimental Results	44
VIII.	RELATED TOPICS TO NOMAC SYSTEM DESIGN	47
	A. The Effect of Nonideal Integration	47
	B. The Effect of Distortion in Multipliers	50
IX.	CONCLUSIONS	53
	REFERENCES	55
	ACKNOWLEDGMENT	56
APPENDIX		
I.	PER-UNIT EQUIVOCATION AND PROBABILITY OF ERROR	57
II.	THE DISTRIBUTION OF VECTOR MAGNITUDES	59
III.	PROBABILITY OF ERROR: ADAPTATION FROM S.O. RICE	63
IV.	THE ARBITRARY ORIENTATION OF THE VECTOR FIELD	65
V.	THE PROBABILITY DENSITY FUNCTION FOR CORRECT OUTPUTS	69
VI.	INTEGRATION TO OBTAIN THE PROBABILITY OF ERROR	73
VII.	SCHEMATICS OF THE EXPERIMENTAL NOMAC SYSTEM	75
VIII.	THE DISTRIBUTION OF SUMS OF PRODUCTS	81

CHAPTER I

THE BASIC COMMUNICATION SYSTEM

A. Elements of Communication Theory

The theory of information has occupied a position of growing importance in communication engineering recently. The main purpose of the theory is to provide means for a quantitative analysis of communication systems. How this purpose is achieved is illustrated by the following discussion of a basic communication system.

The communication process starts with the selection at the transmitter of one member of a set of possible messages, which in the idealized model take the form of symbols. (In electrical communication systems, a "symbol" is usually a voltage or current waveform.) The selection is made at the direction of the originator, and the selected symbol is transformed into a form appropriate for transmission in the channel. The receiver function is to indicate which of the set of possible symbols was selected; if this is done correctly, the communication link has performed its task perfectly (see Fig. 1.1).

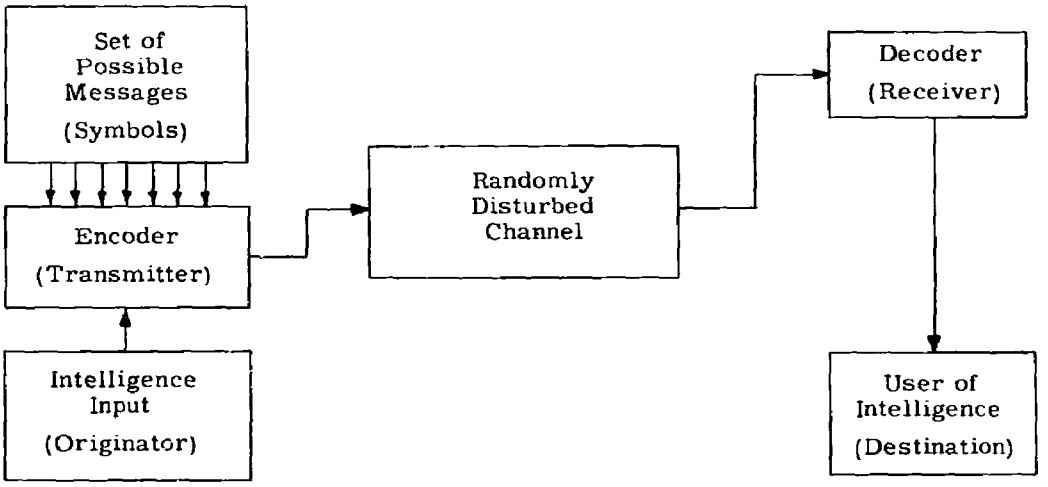

Fig. 1.1. Block diagram of general communication system.

The interpretation of the operation of a communication system as a process of selection of one from a number of possibilities was made by Nyquist in 1924.[17]* This interpretation was later used by Hartley in a 1928 paper[12] in which the logarithm of the number of possible symbols is suggested as a quantitative measure of the information conveyed by the selection.

The introduction of statistical concepts in information theory led to a more general measure of information in terms of the logarithm of the reciprocal of the probability that

*Refer to numbered references at end of report.

a symbol of the set should be selected. (Obviously, where the set consists of equally probable symbols, this measure becomes identical with that suggested by Hartley.) An account of this phase of information theory is contained in the literature, and the reader is particularly directed to the writings of Wiener,[25] Shannon,[22] and Fano.[5]

If noise is introduced in the channel, it has the effect of distorting the received signal in such a way that the distortion might have resulted from more than one of the possible symbols at the transmitter, so far as the receiver is concerned. The receiver, therefore, cannot be certain about what was transmitted because of the uncertainty connected with the disturbing noise, and thus is inherently subject to errors in the decoding process.

From the point of view of information theory, the information about the transmitted symbol at the input of the receiver can be expressed in terms of the change in the probabilities of the possible symbols upon receipt of a signal. Thus, where the set X represents the possible symbols at the transmitter and Z represents the signal at the input of the receiver when one of the X's is selected and transmitted, the information gain associated with each of the symbols is given by

$$I(X/Z) = \log \frac{P(X/Z)}{P(X)} \quad . \tag{1-1}$$

In Eq.(1-1), P(X) is the probability of the symbol while P(X/Z) is the conditional probability of the symbol following the reception of Z. The information received about the transmitted symbol is the average of that shown in Eq.(1-1), averaged over all X that could have resulted in the particular Z received, namely,

$$I(Z) = \sum_X P(X/Z) \log \frac{P(X/Z)}{P(X)} \quad . \tag{1-2}$$

Note that, in the absence of interfering noise, one of the P(X/Z) would be unity, all others, zero. Then Eq.(1-2) reduces to the measure given in the earlier paragraph, i.e., $\log 1/P(X)$, which measures the information associated with the selection made at the transmitter, and thus is the measure of the transmitted information.

The point of view of the discussion of the preceding paragraph is that of Woodward and Davies.[26] They concluded that the best a decoder or receiver can possibly do is to compute the conditional probabilities of the transmitted symbols when a signal appears at the receiver input. To demonstrate how this might be done, Woodward and Davies considered the case where the disturbance in the channel is an independent additive Gaussian white noise. They show that the conditional probabilities are given by a decreasing function of the mean square difference between the received signal and the waveform representing the symbol for which the probability is being computed. Thus, for the symbol X_k,

$$p(X_k/Z) = B\, p(X_k) \exp\left[- \frac{\int_o^T [Z(t) - X_k(t)]^2 \, dt}{N_o} \right] \tag{1-3}$$

Here B is a normalizing constant and N_o is the noise power per cycle of bandwidth. If the X's are equally likely, $p(X_k)$ is a constant for all k, and $p(X_k/Z)$ is a function only of the integral of the squared difference $[Z(t) - X_k(t)]^2$.

If a receiver performs the computation of the conditional probabilities and makes these quantities available to the user of the information, it has relayed all the received information to the output circuit. An examination of Eq.(1-2) reveals that on the average I(Z) is less than the transmitted information when noise is present. The information lost is termed equivocation.

It is also apparent that a receiver that computes the conditional probabilities in the manner described by Eq.(1-3) performs a comparison of the received signal with each of the possible transmitted symbols. It is correctly implied that copies of the possible symbols must be made available at the receiver for the comparison.

One of the problems in communication theory has been how to select the optimum set of signals into which the transmitter encodes the information so as to lead to a minimum amount of equivocation under the condition of fixed rate of transmitted information. This problem has not been solved in general. It has been shown, however, that sets do exist which can lead to ratios of equivocation to transmitted information that are arbitrarily small, if sufficient delay is allowed in the communication process. This is true provided the rate of transmission of information does not exceed the maximum rate at which information may be received through any particular communication channel.

This fact was stated as an existence theorem and proved by Shannon.[21] In a particular form that is applicable to continuously varying time functions disturbed by white Gaussian noise, Shannon derived the maximum rate, called channel capacity, which is given by

$$C = W \log\left(1 + \frac{S}{N}\right) \text{ bits/sec} \quad . \tag{1-4}$$

In this expression, W is the (ideal rectangular) bandwidth occupied by the time-varying signal, S is the component of received signal power due to the transmitted signal, and N is the noise component of the received signal power. An important condition leading to the derivation of Eq.(1-4) is that the average transmitted power is limited.

In his derivation Shannon noted that each of the waveforms of the set of signals that would lead to a full utilization of the system capacity with arbitrarily low equivocation would be in all respects similar to white Gaussian noise. Although unable to determine a particular optimum set of noise-like waveforms, he was able to show that the average performance of all possible noise waveforms (having the same bandwidth and average power) was ideal, in that the fraction of transmitted information lost could be held arbitrarily low at information rates less than the channel capacity, provided sufficient delay were allowed. This performance could be obtained with a receiver that makes a decision about what was transmitted based on the minimum mean square difference between the received signal and each of the possible symbols. A practical and fortunate corollary is that random segments of Gaussian white noise can be used for symbols, and that these segments of noise can be taken from currently generated waveforms from a continuous noise source.

Here a new concept of the role played by the decoder or receiver has been introduced. It has been stated that the best a receiver can do when noise is present is to compute the conditional probabilities of the transmitted symbols following reception of a signal. However, in a practical case, the receiver is usually called upon to indicate which of the set of possible symbols was transmitted. The indication is performed after a decision by the receiver of which

symbol should be indicated. Because of the interfering noise, the receiver cannot decide with certainty, but must choose a symbol that has at least a high probability of having been transmitted. In a typical case, the receiver might indicate the most probable of the transmitted symbols (as Shannon's receiver does).

The indication made by the receiver is, of course, subject to errors. The ratio of the number of erroneous decisions to total decisions is termed the probability of error, and is closely related to the fraction of transmitted information that becomes lost – the per-unit equivocation.*

When the receiver makes a decision about what is transmitted, and only this decision is relayed to the destination, the user of the information knows only that the receiver has selected one of the set of possibilities in accordance with some detection criterion. The information about the conditional probabilities is otherwise discarded, and the over-all per-unit equivocation is thereby increased.

While the per-unit equivocation is the proper criterion for the evaluation of the efficiency with which a communication link performs its assigned task, the probability of error is often used instead. The relation between this relative frequency of erroneous decisions and the per-unit equivocation when the user of information gets only the decisions of the decoding device (receiver) is given in Appendix I. It is shown that the per-unit equivocation is a decreasing function of decreasing probability of error, which justifies the use of the latter in evaluating the performance of the link.

B. The Mathematical Model

The geometrical model employed by Shannon,[12] and which was subsequently used by Rice, was also adopted for the theoretical work in this paper. It is useful in relating the concepts and results found in this research to those found by these two earlier authors; it is outlined briefly here.

In its simplest form, it is assumed that the symbols used are segments of a time function which has (1) a Gaussian amplitude distribution, (2) a flat frequency distribution to W cycles per second with no component frequencies higher than W cycles per second, and (3) an amplitude variance S (which becomes the average power for electrical signals). Each of these segments lasts just T seconds.

It is obvious that, because of the abrupt start and stop of the segments as described, frequency components outside the bandwidth W cannot be avoided. This suggests the characterization of each of the segments in terms of the Fourier coefficients in the manner given by

$$X(t) \doteq \frac{X_o}{\sqrt{2TW}} + \sum_{i=1}^{TW} \frac{X_i}{\sqrt{TW}} \cos \frac{2\pi i}{T} t + \sum_{j=1}^{TW} \frac{X_j}{\sqrt{TW}} \sin \frac{2\pi j}{T} t \quad . \tag{1-5}$$

*For a complete discussion of this relation, see R. M. Fano's printed lecture notes of yearly course, 6.574, Statistical Theory of Information, M. I. T.

Here

$$\frac{X_o}{\sqrt{2TW}} = \frac{1}{T}\int_o^T X(t)\,dt \quad . \tag{1-5a}$$

$$\frac{X_i}{\sqrt{TW}} = \frac{2}{T}\int_o^T X(t)\cos\frac{2\pi i}{T}t\,dt \quad , \tag{1-5b}$$

$$\frac{X_j}{\sqrt{TW}} = \frac{2}{T}\int_o^T X(t)\sin\frac{2\pi j}{T}t\,dt \quad . \tag{1-5c}$$

The X_i or X_j represent the Fourier coefficients, and $X(t)$ is the segment of noise represented. The constant $1/\sqrt{TW}$ is introduced arbitrarily to control the relative magnitudes of the X's. It can be shown that, for the time function as described, the X's are numbers drawn from a normal distribution with variance equal to the power S. By setting an upper limit to the frequencies for which X_i are defined, the higher-frequency components introduced by the starting and stopping of a segment are neglected. As developed, the time functions described will be low-pass functions. However, the description of band-pass functions can be achieved analogously by running the index of summation from $i = TW_1$ to $i = TW_2$ where the frequency band between W_1 and W_2 is the part of the spectrum occupied by the function.

Another description is necessary to support the material presented in Chapter VIII. Although it is somewhat easier to visualize, it is limited in application to low-pass functions. It is also useful in interpreting the vector model introduced later in this chapter. Here, each segment is represented by amplitude samples spaced each $1/2W$ seconds along the waveform. Thus, there are $2TW$ samples for each symbol represented by a finite-duration noise-like waveform. Because the time function has a flat rectangular spectrum, it can easily be shown that the amplitude samples are incoherent; and, because they are from a Gaussian process, they are in fact independent. These samples can thus be used with a sequence of orthonormal functions of the form $\sin t/t$ to reconstruct an approximation to the original waveform. One takes

$$X(t) \doteq \sum_{i=1}^{n} X_i \frac{\sin 2\pi W(t - t_i)}{2\pi W(t - t_i)} \quad , \tag{1-6}$$

in which $t_i = i/2W$. The $X(t)$ resulting from this summation not only differs from the true value in the intersample time regions, but differs from zero outside the duration T. It is, however, a least-mean-square approximation to the true value, as good as can be done with the $2TW$ specifying numbers. It is in that respect entirely equivalent to the representation in terms of Fourier coefficients. Like the Fourier coefficients, it is apparent that the X's in Eq.(1-6) are also numbers chosen from a normal distribution of variance S.

The $n = 2TW$ numbers thus chosen to represent each symbol are ordered and taken as coordinates of a point in Euclidean n-space. Each symbol waveform corresponds uniquely to a point in the space, such that there are $K + 1$ of the points designated by $X_o, X_1 \ldots X_k$. Connecting each point and the origin are $K + 1$ vectors denoted by $\vec{X}_k$. Either a point or a vector may be used to represent the message waveform to which it corresponds.

It is evident that any point in n-space specifies a waveform of the same bandwidth and time duration as the possible message waveforms. The sum of two waveforms, by summing coordinates, becomes the vector sum of the two waveform vectors.

One of these vectors (arbitrarily $\vec{X}_0$) is selected and transmitted. In the channel, an independent white Gaussian noise, $\vec{Y}$, is added. The received signal is by linear superposition the vector sum (see Fig. 1.2) of $\vec{X}_0$ and $\vec{Y}$ and is designated $\vec{Z}$. The receiver used by Shannon and Rice has available copies of the set of vectors $\vec{X}_k$ and chooses, as the one most probably sent, the one whose terminal point is closest to point Z. Here it is evident that the mean square difference between the received signal and each member of the set of possible signals is equal to the squared distance from the point to each point X_k, aside from a constant factor, 1/n. (The term "mean" in mean square difference is used to connote the average over the duration T of the difference $[Z(t) - X(t)]^2$.)

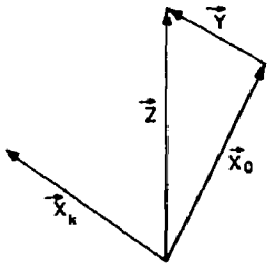

Fig. 1.2 Vector model of messages, noise, and received signal.

Since the interfering noise has been assumed to be white Gaussian, the Woodward and Davies treatment giving the conditional probabilities in terms of the mean square difference is valid. Therefore, it follows that the conditional probability of the k^{th} symbol is a monotonically decreasing function of increasing distance of the point X_k from Z.

The probability of error for a communication system, choosing as the transmitted waveform the one that is the minimum distance from the received signal waveform (in terms of the geometrical model), was the subject of Rice's paper. When modified for small signal-to-noise ratios, in terms of the channel capacity C and information rate at the source $H = \log_2(K + 1)$, the probability of error as obtained in Appendix III is

$$P(\text{error}) \sim \frac{1}{2\sqrt{\pi C'T}} \exp\left[-C'T\left(1 - \frac{H}{C}\right)\right] \quad . \tag{1-7}$$

This result is valid for large n, for large number of possible messages, K + 1, and for a signal-to-noise ratio sufficiently small that the approximation $nS/2N = C'T$ is valid. The prime indicates C' is in logarithmic units. In bits, C is equal to $[\log_2 e]C'$.

It is evident from Eq.(1-7) that the probability of error and thus the per-unit equivocation may be made arbitrarily small with increasing delay T, provided only that H/C, the ratio of transmitted rate of information to system capacity, does not exceed unity.

Rice did not give Eq.(1-7) explicitly in his paper, but gave an expression for the probability of no error, valid for large n and large K, but for all signal-to-noise ratios. He showed that his expression approaches unity under the conditions above. However, his expression was given along with error terms, which decreased with increasing n and K but which exceeded the difference between his expression and unity. Thus, a probability-of-error expression could not be obtained directly from his results, but an intermediate result had to be taken and worked into the form of Eq.(1-7) as is shown in Appendix III. The fact that the expression

given here is limited to cases where the signal-to-noise ratio at the receiver is small is not a serious disadvantage, as will be seen later.

It was mentioned earlier that a receiver may use other criteria in deciding, after a signal has been received, which one of the possible waveforms was transmitted. When such criteria are proposed, they must be evaluated in terms of the probability of error that accompanies the use of that criteria. Of course, other prevailing conditions, such as the method of coding and type of interference, must be given due consideration. In the class of idealized systems in which the possible message waveforms are random samples of white Gaussian noise and the interfering noise is also Gaussian, the rate of errors can be compared with the result obtained by Rice. Rice's results may be regarded as expressing the performance of an ideal system in that his receiver will always select the most probable transmitted symbol (under the conditions of additive Gaussian noise and equal symbol probabilities).

CHAPTER II

CORRELATION DETECTION CRITERIA

Another criterion of detection that is of primary interest in this paper can be obtained as follows. Where the various possible transmitter symbols are given by waveforms $X_k(t)$, $0 < t < T$, and the received signal is $Z(t) = X_o(t) + Y(t)$ (where $Y(t)$ is additive noise), the minimum mean square difference (msd) is given by the minimum of

$$\text{msd} = \frac{1}{T}\int_o^T \left[Z(t) - X_k(t)\right]^2 dt \quad , \tag{2-1}$$

$$= \frac{1}{T}\int_o^T \left[Z(t)^2 + X_k(t)^2 - 2Z(t)X_k(t)\right] dt \quad . \tag{2-2}$$

If it is assumed that

$$\frac{1}{T}\int_o^T X_k(t)^2 dt$$

is a constant for all k, a minimum of Eq. (2-2) is a maximum of the term

$$\frac{1}{T}\int_o^T Z(t)X_k(t)\,dt \quad .$$

This may be recognized as the "correlation coefficient" of X_k and Z, and a receiver that computes this value for subsequent use in deciding what was sent is called a correlation detector.[7]

Correlation techniques were introduced to communication engineers largely by N. Wiener, and the techniques have been improved and implemented by Lee[16] and others.[15] The cross-correlation function of two stationary random functions of time is given by

$$\phi_{12}(\tau) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^T f_1(t) f_2(t+\tau)\,dt \quad . \tag{2-3}$$

When $f_2(t)$ is identically $f_1(t)$, this becomes the autocorrelation function. The relation between autocorrelation functions and power spectra[24] makes correlation techniques an invaluable aid in the study of random functions.

The so-called "short-time cross-correlation function" has been studied[6] and is given by

$$\phi_T(\tau) = \frac{1}{T}\int_o^T f_1(t) f(t+\tau)\,dt \quad , \tag{2-4}$$

which is seen to be of the form of the defining expression for correlation coefficients. A generalized form of Eq. (2-4) in which the product $f_1(t) f_2(t+\tau)$ is filtered rather than mathematically integrated is

$$\phi(t,\tau) = \int_{-\infty}^{\infty} h(t-\sigma) f_1(\sigma) f_2(\sigma+\tau)\,d\sigma \quad . \tag{2-5}$$

Here h(t) is the impulse response function of the filter and is sometimes variously called the integrating function, scanning function, or window function.[24]

The use of correlation devices operating as practical correlation detectors in the manner described mathematically by Eq.(2-5) has brought about the adoption of detection criteria based on correlation outputs in their own right. An advantage of such correlation devices is that they are not subject to an important limitation on a device computing the mean square difference. The latter device must know the precise value of the component of received signal power due to the transmitted waveform as well as the waveform shapes themselves. However, the former device need know the waveforms only within the freedom allowed by an arbitrary constant multiplier.

In reference to Eq.(2-2), it is apparent that if the integrals

$$\int_o^T X_k(t)^2 \, dt$$

(proportional to the signal energy) are not equal for all k, the criterion of maximum short-time correlation is not equivalent to that of minimum mean square difference. To illustrate this, the geometrical model is consulted.

The additive properties of the vector representations of the time functions were shown to follow from linear superposition. The orthogonality of the components leads to equivalent vector interpretations of the average product or correlation coefficient. For example, dealing only with the cosine terms of the representation given in Eq.(1-5),

$$\frac{1}{T}\int_o^T X(t)Z(t)\,dt \doteq \frac{1}{T}\int_o^T \sum_{i=1}^{TW} \frac{X_i}{\sqrt{TW}} \cos\frac{2\pi i}{T}t \; \Sigma \; \frac{Z_k}{\sqrt{TW}} \cos\frac{2\pi k}{T}t\,dt + (\text{other terms}), \qquad (2\text{-}5)$$

$$\doteq \sum_{i=1}^{TW}\sum_{k=1}^{TW} \frac{X_i Z_k}{TW} \left\{ \frac{1}{T}\int_o^T \cos\frac{2\pi i}{T}t \cos\frac{2\pi k}{T}t\,dt \right\} + (\quad), \qquad (2\text{-}5a)$$

when the order of integration and summation is interchanged. But

$$\frac{1}{T}\int_o^T \cos\frac{2\pi i}{T}t \cos\frac{2\pi k}{T}\,dt = \frac{1}{2}\delta_i^k \qquad (2\text{-}6)$$

where δ_i^k is the Kronecker delta and $\delta_i^i = 1$; $\delta_i^k = 0$, $i \neq k$.
Also, the same behavior governs the sine terms, and all the cross terms are zero. Finally,

$$\frac{1}{T}\int_o^T X(t)Z(t)\,dt = \sum_{i=1}^{TW} \frac{X_i Z_i}{2TW} \qquad (2\text{-}7)$$

(The summation on the right is obtained approximately when the alternate representation of the symbol waveforms in terms of 2TW time samples is employed.)

The $\sum_i X_i Z_i$ is obviously the form of the dot or scalar product of the vectors $\vec{X}$ and $\vec{Z}$. It is related to the distance between X and Z by

$$\vec{X}_k \cdot \vec{Z} = \frac{1}{2}|Z|^2 \left\{ 1 + \frac{|X_k|^2}{|Z|^2} - \frac{D_k^2}{|Z|^2} \right\}, \qquad (2\text{-}8)$$

where D_k is used to denote the distance corresponding to the k^{th} vector. Again, it is evident that if all $|X_k|^2$ are equal, the minimum distance corresponds to the maximum dot product. As is shown in Appendix II, however, the magnitudes of the vectors $\vec{X}_k$ are not all equal when they are segments taken from a random Gaussian noise waveform. Where the signal power S is large compared to the noise power N, a case such as that shown in Fig. 2.1 always has finite probability. Here, since $|\vec{Z}|$ is constant for all k, the correlation is that constant times the projection of the vector $\vec{X}_k$ onto the vector $\vec{Z}$. It is immediately evident that a vector might occur with magnitude sufficient that, although the distance $D_k > D_o$, the projection of $\vec{X}_k$ will exceed that of $\vec{X}_0$ on $\vec{Z}$.

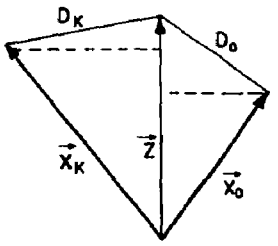

Fig. 2.1. Vector magnitude contributions of error.

When the ratio of signal power S to noise power N becomes quite small, the term $|X_k/Z|^2$ itself becomes small. Then the contribution of the fluctuations of the magnitude of $\vec{X}_k$ to the values of the dot products is to the value of $\vec{X}_k \cdot \vec{Z}$ approximately as the signal amplitude is to the noise amplitude. Thus, for small signal-to-noise ratios, the equivalence of the criteria of maximum dot product and minimum distance is again approached, even though the energies of each $\vec{X}_k$ are not equal. As is seen later (for all S/N), as the system dimension n is increased, the fluctuations of $|X_k|^2$ decrease percentagewise, further establishing the equivalence of these two criteria.

The fact that the criterion of maximum correlation is an optimum one for small signal-to-noise ratios, $\rho = S/N$, is not of mere academic interest. It is when ρ is small that the communication of information becomes most difficult. Here the channel capacities for channels of bandwidth conventionally associated with communications become low enough that the information rates of even relatively low-rate systems, such as telegraph and teletype, become a significant fraction of the channel capacity.

Furthermore, there are two features of military significance that are inherently related to low signal-to-noise ratios.

One of these is the possibility of communicating with received signal levels below the receiver and antenna noise at the receiver location. In the past, other systems have been proposed which provide communication, although the average signal power is less than the noise power at the receiver input. However, these systems, such as pulse-position or pulse-code systems, feature bursts of power for relatively short periods of time which are above the noise level. The average power is less than the noise power by virtue of its being averaged over larger periods of time. In the proposed systems, the signal power can be less than the noise power at all times. If unfriendly search receivers are limited to a comparable received power, it is highly unlikely that such receivers will be aware of the presence of the signal "on the air," unless, of course, the unfriendly receiver performs the same operation of correlation as the friendly one.

Another promising feature deals with resistance to jamming. As Fano has shown,[7] the signal-to-noise ratio at the output of a correlation detector is given ideally by the ratio of total signal energy to noise power per cycle. This suggests the jamming power may be forced

down to a low value compared with the signal energy merely by spreading the signal energy over a sufficiently wide bandwidth. This spread causes the jamming power to be effectively averaged over the bandwidth in such a way that the jamming power per cycle is small.

In the next chapter, a specific description of the systems suggested by the properties of this type of noise communication and correlation is accompanied by a more complete discussion of what services these systems are expected to perform and their advantage in performing them.

CHAPTER III

DESCRIPTION OF THE FUNDAMENTAL NOMAC SYSTEM

In view of the advantages that schemes of communication using noise-like signals and correlation detectors appear to possess, an extensive investigation into their properties has been conducted. The code word NOMAC (coined from Noise Modulation and Correlation) has been suggested for use in referring to such systems.

NOMAC systems, generally speaking, trade bandwidth for the ability to operate at low signal-to-noise ratios. In view of the present day emphasis on conservation of bandwidth, these systems should not find application where interference is slight or negligible. However, in certain military applications the possibility of maintaining communications secure from intercept or reliable in the face of enemy jamming offsets any disadvantage that may be connected with the use of wide bandwidths.

The block diagram of a fundamental NOMAC system is shown in Fig. 3.1. The set of possible symbols which take the form of finite duration segments of Gaussian noise is shown at the transmitter. The transmitter selects one of these waveforms $X_k(t)$ and propagates it through the channel in which the Gaussian interfering noise $Y(t)$ is added. At the receiver, copies of the K + 1 waveforms are available to use in K + 1 correlation detectors in which the received signal is compared with each of the copies representing possible symbols.

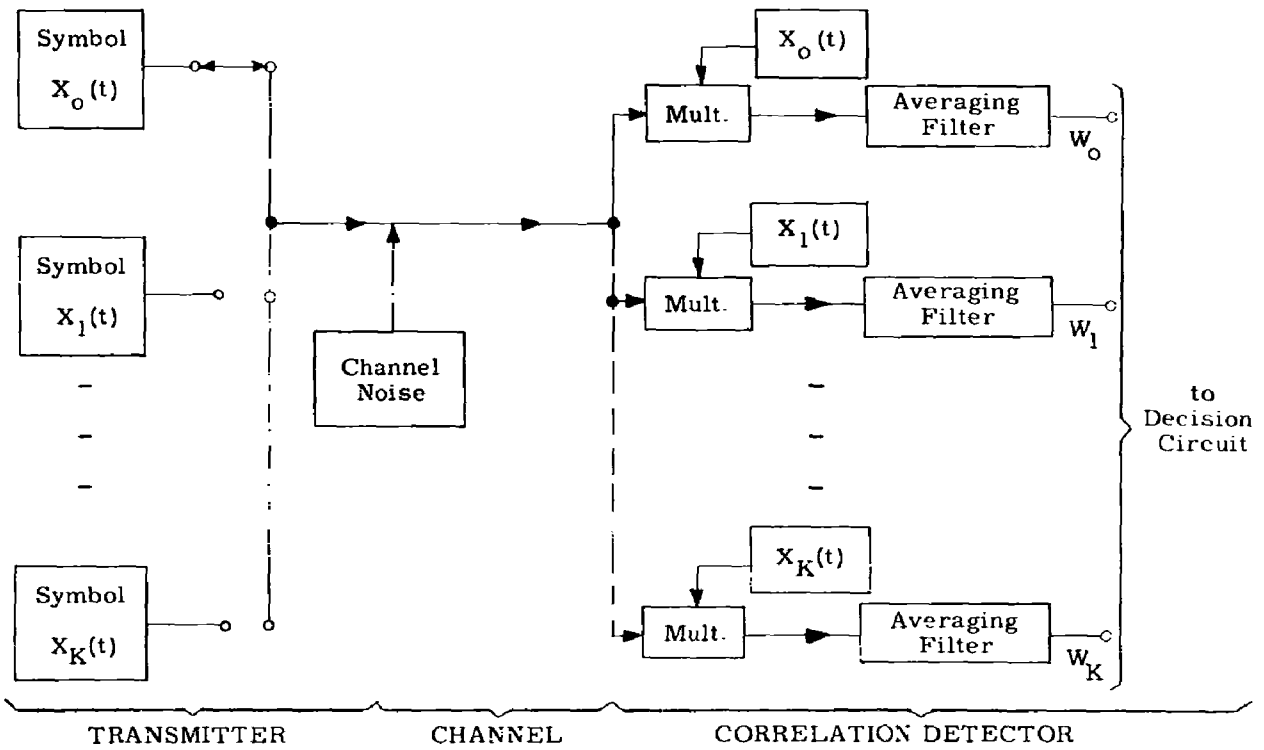

Fig. 3.1. Block diagram of typical NOMAC system.

A problem of major importance in the design of NOMAC systems is that of delivering the copies of the possible waveforms to the receiver so that correlation may be performed. The attempts at a solution of this problem divide NOMAC systems into two categories, one in which the possible symbol waveforms are stored at the receiver, and another in which the waveforms reach the receiver as reference signals through one or more auxiliary channels.

The first category, called the stored-signal system, presents rather severe requirements of time synchronization, in order that the correlation coefficient calculated will correspond to the $\tau = 0$ point of the autocorrelation function. One advantage of this system is that, while the segments may be chosen at random from white Gaussian noise, they become a known set once the choice is made. Thus, a scale factor may be employed to make each of the symbol energies equal, which establishes the exact equivalence of the criteria of maximum correlation and of minimum mean square difference.

In the second category, the synchronization problem is largely eliminated, but noise is generally present in the auxiliary channels also. Here the signals may be randomly selected from one or more noise sources that are currently generating the noise. Obviously, these symbol waveforms will be random in all respects.

In Fig. 3.2, curves of the signal-to-noise ratio in the output of an ideal correlation detector as a function of the input signal-to-noise ratio and $n = 2TW$ are shown for the two categories of NOMAC system. It will be shown subsequently that, when filtering is used for the integration in the correlation process, n is in reality the ratio of the input-signal bandwidth W to the noise bandwidth of the integrating filter. In the figure, the signal-to-noise ratio ρ, or $(S/N)_c$, is assumed the same for both intelligence and auxiliary channel or channels. From the figure, if a required output ratio and n are known, one may determine the permissible input signal-to-noise ratio. Conversely, if a desired input signal-to-noise ratio and required output signal-to-noise ratio are specified, the necessary bandwidth ratio is easily obtained for either type of system.

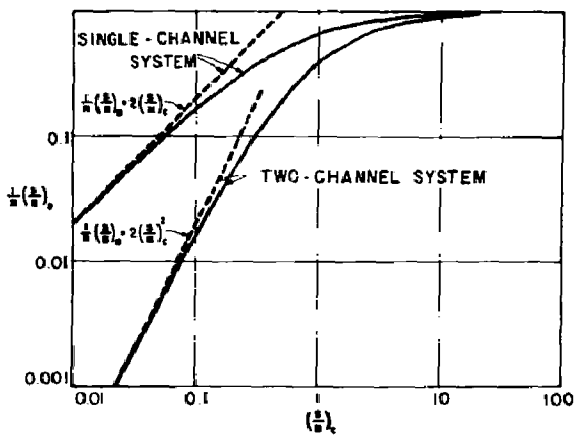

Fig. 3.2. Signal-to-noise improvement in correlation vectors.

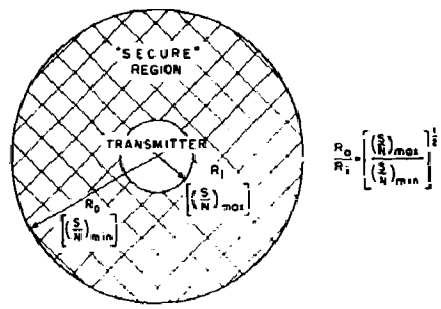

Fig. 3.3. Secure region of communication.

To demonstrate the secure communication properties of such a system, suppose that the NOMAC system as designed will operate satisfactorily at an input signal-to-noise ratio $(S/N)_{min}$, while an unfriendly search receiver can detect the radiation only at signal-to-noise ratios greater than a larger value $(S/N)_{max}$. Then, as is seen in Fig. 3.3, there is a "secure" region or annular ring about the transmitter in which communication might be carried out by friendly parties without the knowledge by unfriendly forces in that region.

An inverse line of reasoning would demonstrate that an unfriendly jamming transmitter might be believed to be effective against the assumed transmitted power as determined by the power received at the jamming site to some range R_o. On the other hand, communication would actually be maintained to the inner radius R_i, and the "secure region" now becomes a marginal communication area which might conceivably prove quite embarrassing to the unfriendly forces.

As presented in Fig. 1 the fundamental NOMAC system is idealized to facilitate the theoretical investigation. The type of decision circuit is not specified in order to permit some latitude in the interpretation and application of the figure. In the following material, two types of decision circuit are evaluated in detail. The first of these is one that establishes the criterion of maximum correlation in its selection of which of the symbols X_k was transmitted. The second decision circuit (and the easiest to construct in practice) sets as a criterion the exceeding of a fixed threshold. Called threshold detection, it indicates a signal as having been transmitted whenever the correlator output corresponding to that signal exceeds the threshold value.

It is not intended to convey the impression that only systems designed precisely as indicated by the block diagram are included in the analysis presented. Modifications of the analysis given here for discrete systems, along with the work appearing elsewhere concerning the improvement of signal-to-noise ratio in correlation detectors,[7] can be made to extend the coverage to a varied class of similar systems. For example, one might use a single random-noise source to obtain the different symbols merely by using delayed versions of the initial noise for the sources of the currently chosen segments. The delay increments need only be great enough to correspond to values of τ of the autocorrelation function of the noise for which the correlation function is essentially zero.

Other versions of NOMAC systems may include those in which the noise-like waveform is modulated in the same manner in which a sinusoidal carrier is modulated in conventional communication systems. For example, the transmitted random waveform might be varied in amplitude, frequency band of transmission, or relative time of transmission. These variations correspond to conventional AM, FM and PPM, for example. Obviously, combinations of any of these modulations are possible just as with sinusoidal carriers. The experimental model shown in block diagram in Appendix VII, Fig. 3 is essentially an amplitude-modulated version of a NOMAC system, and its theoretical probability-of-error treatment corresponds to the signal-to-noise improvement type of analysis that would be used for an AM system.

CHAPTER IV

THE THEORETICAL STUDY OF PROBABILITY OF ERROR

A. The Criterion of Maximum Correlation

When the receiver of a NOMAC system makes its decision about which of the signals $X_K(t)$ was transmitted – based on the criterion of maximum correlation output – the rate of making errors has been obtained. The result, which is valid for large n, and for signal-to-noise ratio ρ, such that the product $n\rho$ is larger than about 2, is given by the approximation

$$P(\text{error}) \sim \frac{K}{2} \operatorname{erf} \sqrt{\frac{n\rho}{1+\rho}} \quad , \tag{4-1}$$

where

$$\operatorname{erf} Q = \sqrt{\frac{2}{\pi}} \int_Q^\infty \exp\left[-\frac{1}{2}t^2\right] dt \quad .$$

For still larger values of $n\rho$, but only for small signal-to-noise ratios, the expression agrees with that of Rice, namely,

$$P(\text{error}) \sim \frac{1}{2\sqrt{\pi C'T}} \exp\left[-C'T\left(1-\frac{H}{C}\right)\right] \quad . \tag{4-1a}$$

Here, $\rho \ll 1$, so that $\frac{1}{2}n\rho \doteq C'T$; and $K \gg 1$, so that $\ln K \doteq \ln K + 1 = H'T$.

These results were obtained using the geometrical model in the following manner. First, there are K + 1 independent message vectors which are represented by $\vec{X}_0, \vec{X}_1, \ldots \vec{X}_K$. Appendix II derives the probability density distribution for the magnitude of these vectors, a type familiar in the study of statistics For the message vector $\vec{X}$, where S is the average signal power, it is given as follows:

$$p_o(|X_i|) = \frac{2|X_i|^{n-1}}{(2S)^{n/2}\Gamma(\frac{n}{2})} \exp\left[-\frac{|X_i|^2}{2S}\right] \quad . \tag{4-2}$$

As is shown in Appendix II, this distribution has an average value of

$$\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\sqrt{2S}$$

and a variance of not more than 1/2S. The average value is $\sqrt{nS}$ within an error of one part in each 4n parts.

Here it must be pointed out that approximations are made throughout the paper, primarily because the representation of the actual waveforms as having a Gaussian probability density distribution is only approximately correct,[14] and it may be considerably in error along the "skirts" of the distribution curve. Except where otherwise noted, linear systems are assumed, which is not necessarily the case in practice and is certainly not true over the whole range of amplitude values for which the Gaussian distribution is defined. Since all the results

obtained depend on these two assumptions (and other assumptions), results may in some cases be true only to the order of magnitude represented. However, the primary purpose, which is that of indicating the behavior one might expect in a physical system and of demonstrating how the behavior varies with important parameters, is served.

In Fig. 4.1, the combination of one of the message vectors, say, $\vec{X}_0$, and the interfering noise $\vec{Y}$ are shown resulting in the vector $\vec{Z}$. Also, another typical message vector $\vec{X}_k$ is shown. The points X lie approximately on a hypersphere of radius $\sqrt{nS}$, while Z lies approximately on a hypersphere of radius $\sqrt{nP}$. Here P is the total power S + N where S is the signal power and N is the interfering noise power. The NOMAC receiver cross-correlates the received signal, vector $\vec{Z}$, and each of the message waveforms, vectors $\vec{X}_k$. As stated before, this is equivalent to taking the dot product between the $\vec{Z}$ vector and each of the set of message vectors aside from the constant $1/n$. The resulting set of dot products is designated by W_k.

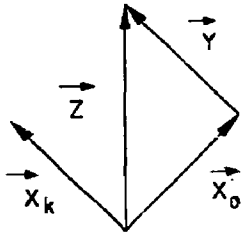

Fig. 4.1. Vector model of messages, noise, and received signal (Fig. 1.2).

Up to this point, the analysis of a NOMAC system does not depend on how the receiver makes use of the outputs W_k. However, for further study it is necessary to distinguish between the two cases: that in which the receiver decision circuit indicates the largest W_k as the one determining the transmitted signal; and that in which a W_k that exceeds a previously established threshold value is the indication of the transmitted signal.

For the case in which the criterion of detection is that of maximum correlator output, it is evident that no error will occur if the dot product W_0 exceeds all other W_k. This in turn is the case if the component of the vector $\vec{X}_0$ along $\vec{Z}$ is greater than the Z-components of all other vectors $\vec{X}_k$. Since the coordinates of the vectors are independent random values, the vectors themselves are randomly distributed in n-space. Appendix IV illustrates this point and shows that the Z-component of any of the K vectors not selected at the transmitter is chosen from the same density distribution as are the coordinates of the message vectors. Therefore, the probability density distribution of the Z-components is given by the Gaussian function

$$p(X') = \frac{1}{\sqrt{2\pi S}} \exp\left[-\frac{X'^2}{2S}\right] . \qquad (4\text{-}3)$$

The prime is here used to indicate the component in the direction of $\vec{Z}$.

The probability that any one Z-component will be less than the Z-component of $\vec{X}_0$ (hereafter designated X_z) is given by

$$P(X'<X_z) = \int_{-\infty}^{X_z} p(X')\,dX' , \qquad (4\text{-}4)$$

$$= 1 - \int_{X_z}^{\infty} p(X')\,dX' . \qquad (4\text{-}4a)$$

The probability that all K vectors not selected for transmission will have

Z-components less than X_z is given by

$$P(\text{all } X'<X_z) = \left[1 - \int_{X_z}^{\infty} p(X')\,dX'\right]^K \quad . \tag{4-5}$$

The probability of no error is the average of the probability given over all values assumed by X_z. This is indicated formally as

$$P(\text{no error}) = \int_{-\infty}^{\infty} p_1(X_z)\,P(\text{all } X'<X_z)\,dX_z \tag{4-6}$$

from which

$$P(\text{error}) = 1 - \int_{-\infty}^{\infty} p_1(X_z)\left[1 - \int_{X_z}^{\infty} p(X')\,dX'\right]^K dX_z \quad . \tag{4-7}$$

The function $p_1(X_z)$, the probability density distribution function of Z-components of the vector $\vec{X}_0$, must be determined in order to carry out the indicated integration. This is one of the more difficult aspects of the problem, and a considerable portion of the total research has been devoted to finding the properties of and suitable approximations for the density distribution.

B. The Density Distribution of X_z

As shown in Fig. 4.2, the vector $\vec{Y}$ can be broken into two components, one of which is along $\vec{X}_0$ and one normal to $\vec{X}_0$. Furthermore, these two components are independent. This conclusion is reached after noting that the n components making up vector $\vec{Y}$ are independent, and, according to the methods of Appendix IV, may be thought of as independent after rotation of the coordinate axes, since $\vec{Y}$ is independent of the vector $\vec{X}_0$ onto which the initial axis is aligned in this example. There are thus three independent quantities, $|X_o|$, Y_1 and Y_2, that combine to yield vector $\vec{Z}$ in a right-triangle relationship. As expressed, Y_1 is always positive (a magnitude) while Y_2 may be positive or negative.

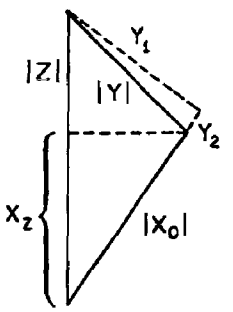

Fig. 4.2. Message, noise, and signal vectors showing components of vectors.

$$|Z| = \sqrt{(|X_o| + Y_2)^2 + Y_1^2} \quad . \tag{4-8}$$

A knowledge of the right triangle formed by these components leads to a fairly simple expression for X_z in terms of these quantities, namely,

$$X_z = |X_o| \frac{|X_o| + Y_2}{\sqrt{(|X_o| + Y_2)^2 + Y_1^2}} \quad . \tag{4-9}$$

A somewhat more easily handled expression is

$$X_z^2 = \frac{|X_o|^2}{1 + \frac{Y_1^2}{(|X_o| + Y_2)^2}} \quad , \tag{4-10}$$

or

$$X_z = \pm \sqrt{\frac{|X_o|^2}{1 + \frac{Y_1^2}{(|X_o| + Y_2)^2}}} \quad \begin{cases} X_z < 0 \text{ if } Y_2 < -|X_o| \\ X_z > 0 \text{ if } Y_2 > -|X_o| \end{cases} \quad . \tag{4-11}$$

Then, by steps,

$$P\left[X_z < a\right] = P\left[-\sqrt{\frac{|X_o|^2}{1 + \frac{Y_1^2}{(|X_o| + Y_2)^2}}} < a\right], \quad a < 0; \text{ (for } Y_2 < -|X_o|) \quad , \tag{4-12}$$

$$= P\left[+\sqrt{\frac{|X_o|^2}{1 + \frac{Y_1^2}{(|X_o| + Y_2)^2}}} < a\right], \quad a > 0; \text{ (for } Y_2 > -|X_o|) \quad . \tag{4-12a}$$

With due regard for the sign of a, an intermediate step follows, valid for both Eq.(4-12) and Eq.(4-12a).

$$P\left[X_z < a\right] = P\left[(|X_o| + Y_2)^2 < \frac{a^2 Y_1^2}{|X_o|^2 - a^2}\right] \quad . \tag{4-12b}$$

Then

$$P\left[X_z < a\right] = P\left[(|X_o| + Y_2) < +\sqrt{\frac{a^2 Y_1^2}{|X_o|^2 - a^2}}\right], \quad \text{for } Y_2 > -|X_o| \quad , \tag{4-13}$$

$$= P\left[(|X_o| + Y_2) < -\sqrt{\frac{a^2 Y_1^2}{|X_o|^2 - a^2}}\right], \quad \text{for } Y_2 < -|X_o| \quad , \tag{4-13a}$$

$$= P\left[(|X_o| + Y_2) < a\sqrt{\frac{Y_1^2}{|X_o|^2 - a^2}}\right], \pm \text{ by sign of a, for } |a| \leq |X_o| \quad , \tag{4-13b}$$

$$= P\left[Y_2 < \frac{aY_1}{\sqrt{|X_o|^2 - a^2}} - |X_o|\right] , \quad \text{for } |a| \leq |X_o| \tag{4-14}$$

Obviously, $|X_z| \leq |X_o|$, and therefore $P(X_z < a)$, is unity for all $a > |X_o|$, while it is zero if $a < -|X_o|$.

Expression (4-14) is derived for fixed $|X_o|$ and Y_1. The average probability that X_z is less than a is obtained by averaging over all $|X_o|$ and Y_1.

$$P_1[X_z < a] = \int_o^\infty p_o(|X_o|)\, d|X_o| \int_o^\infty p_q(Y_1)\, dY_1\, P\left[Y_2 < a\sqrt{\frac{Y_1^2}{|X_o|^2 - a^2}} - |X_o|\right] , \tag{4-15}$$

in which

$$P\left[Y_2 < a\sqrt{\frac{Y_1^2}{|X_o|^2 - a^2}} - |X_o|\right] = \begin{cases} 0, & \text{for } a < -|X_o| . \\ \displaystyle\int_{-\infty}^{\frac{aY_1}{\sqrt{|X_o|^2 - a^2}} - |X_o|} p(Y_2)\, dY_2 \equiv \int_{-\infty}^{\frac{aY_1}{\sqrt{|X_o|^2 - a^2}} - |X_o|} \frac{1}{\sqrt{2\pi N}} \exp\left(-\frac{Y_2^2}{2N}\right) dY_2 , & \\ \text{for } -|X_o| < a < |X_o| . & \\ 1 \qquad \text{for } a > |X_o| . \end{cases} \tag{4-16}$$

A sketch of cumulative probability distribution $P[X_z < a]$ for particular $|X_o|$ and Y_1 is shown in Fig. 4.3.

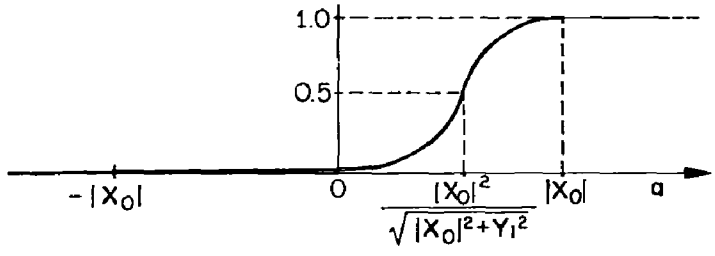

Fig. 4.3. Probability distribution of components.

With the aid of Fig. 4.3, it is seen that, when a is positive,

$$P[X_z < a] = \int_o^a p_o(|X_o|)\, d|X_o| + \int_a^\infty p_o(|X_o|)\, d|X_o| \int_o^\infty p_a(Y_1)\, dY_1 \int_{-\infty}^{\frac{aY_1}{\sqrt{|X_o|^2 - a^2}} - |X_o|} p(Y_2)\, dY_2 \quad ; \tag{4-17}$$

for $a < o$,

$$P[X_z < a] = \int_{|a|}^{\infty} p_o(|X_o|)d|X_o| \int_o^{\infty} p_q(Y_1)\,dY_1 \int_{-\infty}^{\frac{aY_1}{\sqrt{|X_o|^2-a^2}} - |X_o|} p(Y_2)\,dY_2 \quad .$$

The probability density distribution function is obtained by differentiating Eq.(4-17) and Eq.(4-17a) with respect to a (see Ref. 8, page 55).

$$p_1(a) = \frac{d}{da}\int_o^a p_o(|X_o|)d|X_o| + \int_a^{\infty} p_o(|X_o|)d|X_o| \int_o^{\infty} p_q(Y_1)dY_1 \frac{d}{da}\int_{-\infty}^{\frac{aY_1}{\sqrt{|X_o|^2-a^2}} - |X_o|} p(Y_2)\,dY_2$$

$$- \left[p_o(|X_o|) \int_o^{\infty} p_q(Y_1)\,dY_1 \int_{-\infty}^{\frac{aY_1}{\sqrt{|X_o|^2-a^2}} - |X_o|} p(Y_2)\,dY_2 \right] |X_o| = a \quad , \tag{4-18}$$

or

$$p_1(a) = \int_a^{\infty} p_o(|X_o|)\,d|X_o| \int_o^{\infty} p_q(Y_1)\,dY_1\; p\left\{ \frac{aY_1}{\sqrt{|X_o|^2-a^2}} - |X_o| \right\} \frac{d}{da}\left\{ \frac{aY_1}{\sqrt{|X_o|^2-a^2}} \right\}, \tag{4-18a}$$

which becomes

$$p_1(a) = \int_a^{\infty} p_o(|X_o|)d|X_o| \int_o^{\infty} p_q(Y_1)dY_1 \frac{|X_o|^2 Y_1}{(|X_o|^2-a^2)^{3/2}}\, p\left\{ \frac{aY_1}{\sqrt{|X_o|^2-a^2}} - |X_o| \right\} \quad . \tag{4-18b}$$

When a is negative,

$$p_1(a) = \int_a^{\infty} p_o(|X_o|)\,d|X_o| \int_o^{\infty} p_q(Y_1)dY_1 \frac{d}{da}\int_{-\infty}^{\frac{aY_1}{\sqrt{|X_o|^2-a^2}} - |X_o|} p(Y_2)\,dY_2$$

$$- \left[p_o(|X_o|) \int_o^{\infty} p_q(Y_1)dY_1 \int_{-\infty}^{\frac{aY_1}{\sqrt{|X_o|^2-a^2}} - |X_o|} p(Y_2)\,dY_2 \right] |X_o| = -a \quad , \tag{4-19}$$

or

$$p_1(a) = \int_{-a}^{\infty} p_o(|X_o|)d|X_o| \int_o^{\infty} p_q(Y_1)dY_1 \frac{|X_o|^2 Y_1}{(|X_o|^2-a^2)^{3/2}}\, p\left\{ \frac{aY_1}{\sqrt{|X_o|^2-a^2}} - |X_o| \right\} \quad . \tag{4-19a}$$

When written in terms of X_z, the desired probability density distribution function for Z-components of the vectors $|X_o|$ reads

$$p(X_z) = \int_{|X_z|}^{\infty} p_o(|X_o|)d|X_o| \int_0^{\infty} p_q(Y_1)\,dY_1 \frac{|X_o|^2 Y_1}{(|X_o|^2 - X_z^2)^{3/2}}\, p\left\{\frac{X_z Y_1}{\sqrt{|X_o|^2 - X_z^2}} - |X_o|\right\} \quad . \tag{4-20}$$

In Eq.(4-20), the p without a subscript denotes a normal probability distribution function with variance N, while p_o is used to designate a distribution of the type considered in Appendix II, in this case of dimension n derived from a variance S. The other function distinguished by the subscript q is of the same type as p_o, but of dimension n – 1 and with S replaced by N, the noise power.

It is therefore apparent that $p_1(X_z)$ in Eq.(4-20) is a Gaussian distribution function modified by the ratio

$$\frac{|X_o|^2 Y_1}{(|X_o|^2 - X_z^2)^{3/2}} \left(\text{or } \frac{Y_1}{|X_o|}\left[1 - \frac{X_z^2}{|X_o|^2}\right]^{-3/2}\right)$$

and subsequently averaged over all $|X_o|$ greater than X_z and over all Y_1. It is immediately evident that the work necessary to determine a precise expression for $p_1(X_z)$ represents an unwise investment in time, since the precision of the original assumptions does not justify the labor. However, an examination of the factors contributing to $p_1(X_z)$ can lead to a useful approximation.

First, since $p_1(X_z)$ is largely an averaged form of a Gaussian distribution function, it should itself tend to a bell-shaped distribution, particularly for large n for which the averaging probability functions approach impulse functions. It follows that the average value and the variance of the distribution can be used to obtain a suitable approximation under certain conditions.

The average value of X_z is difficult to obtain, but it approaches $\sqrt{n/P}\,S$ for large n and does not differ greatly from this value. This fact is demonstrated by expanding the expression for X_z in terms of $|X_o|$, Y_1, and Y_2 in a Taylor series about the value of X_z assumed when each of these variables takes on its average value. It then can be shown that the expectation $E(X_z - \bar{X}_z)^2$ is an order of magnitude smaller than $|\bar{X}_z|^2$ which represents this average value squared. The first term of the expansion is

$$\Delta X_z = \bar{X}_z - X_z \doteq \frac{\partial X_z}{\partial |X_o|}\,\Delta|X_o| + \frac{\partial X_z}{\partial Y_1}\,\Delta Y_1 + \frac{\partial X_z}{\partial Y_2}\,\Delta Y_2 \quad . \tag{4-21}$$

Since the standard deviations of $|X_o|$, Y_1 and Y_2 are small compared to the vectors making up the triangle (for large n), the first term is used to represent all significant contributions to X_z. From their independence, the variance of the sum is the sum of the variances of the components, namely,

$$\sigma_{X_z}^2 = \left[\frac{\partial X_z}{\partial |X_o|}\bigg|_{\bar{X}_z} \sigma_{|X_o|}\right]^2 + \left[\frac{\partial X_z}{\partial Y_1}\bigg|_{\bar{X}_z} \sigma_{Y_1}\right]^2 + \left[\frac{\partial X_z}{\partial Y_2}\bigg|_{\bar{X}_z} \sigma_{Y_2}\right]^2 \quad . \tag{4-22}$$

$$\frac{\partial X_z}{\partial |X_o|} = \frac{\left[(|X_o| + Y_2)^2 + Y_1^2\right](2|X_o| + Y_2) - (|X_o| + Y_2)^2|X_o|}{\left[(|X_o| + Y_2)^2 + Y_1^2\right]^{3/2}} \quad . \tag{4-23}$$

$$\frac{\partial X_z}{\partial Y_1} = \frac{\left[(|X_o| + Y_2)^2 + Y_1^2\right]|X_o| - (|X_o| + Y_2)^2|X_o|}{\left[(|X_o| + Y_2)^2 + Y_1^2\right]^{3/2}} \quad . \tag{4-24}$$

$$\frac{\partial X_z}{\partial Y_2} = \frac{|X_o| Y_1 (|X_o| + Y_2)}{\left[(|X_o| + Y_2)^2 + Y_1^2\right]^{3/2}} \quad . \tag{4-25}$$

When the partials are evaluated at the mean $\overline{X}_z$ and inserted in Eq.(4-22) the result is

$$\sigma_{X_z}^2 = \frac{S^2\left[S + (2 - \frac{1}{n})N\right]^2 + 2SN^3(1 - \frac{1}{n})^2 + S^2N^2(1 - \frac{1}{n})}{2(S + N)^3} \tag{4-26}$$

Since large n has already been postulated, neglect of the terms $1/n$ puts Eq.(4-26) into the form

$$\sigma_{X_z}^2 = \frac{S}{2}\left[1 + \frac{N}{S + N}\right] \quad . \tag{4-26a}$$

The validity of Eq.(4-26a) depends upon the smallness of the ratio $\sigma_{X_z}/\overline{X}_z$, which is of the order $\sqrt{S} \div \sqrt{n/P}\,S$. Thus, if $nS/P >> 1$, Eq.(4-26a) is useful. The earlier stipulation of large n is consistent with this requirement. It is noted that, for small signal-to-noise ratios, $nS/P \doteq nS/N$. This in turn is

$$\frac{nS}{N} = \frac{2TWS}{N} = 2\frac{E_s}{N_o} = n\rho \quad . \tag{4-27}$$

Here E_s is the total signal energy, N_o the noise power per cycle. Here again, ρ is used to indicate the signal-to-noise ratio in the signal or intelligence channel and $2(E_s/N_o)$ and $n\rho$ are used interchangeably, the latter being preferred when brevity is an asset.

The average value and the variance lead to the writing of an approximation for $p_1(X_z)$ in Gaussian form, which is valid for large n as follows:

$$p_1(X_z) \doteq \frac{1}{\sqrt{2\pi\sigma_{X_z}^2}} \exp\left[(-1/2\,\sigma_{X_z}^2)\,(X_z - S\sqrt{\frac{n}{P}})^2\right] \quad . \tag{4-28}$$

This expression is reasonably close to the true distribution over the bell-shaped part of the curve, which is the significant portion of the distribution. Fortunately, while Eq.(4-28) may be considerably in error along the skirts, its role in the integration is only one of averaging, and the parts most in error contribute least to the final result. The related distribution

of the normalized value $X_z \div S\sqrt{n/P}$ is

$$p_1'\left(\frac{X_z}{\sqrt{\frac{n}{P}}S}\right) = \frac{1}{\sqrt{\pi\left(\frac{S+2N}{nS}\right)}} \exp\left[-\frac{nS}{(S+2N)}\left(1-\frac{X_z}{\sqrt{\frac{n}{P}}S}\right)^2\right] \quad . \tag{4-28a}$$

When this substitution is properly made in Eq. (4-7), the result obtained is

$$P(\text{error}) \doteq 1 - \int_{-\infty}^{\infty} \frac{\exp\left[-\frac{nS}{S+2N}(1-X)^2\right]}{\sqrt{\pi\frac{(S+2N)}{nS}}} \left[1 - \int_{\sqrt{\frac{nS}{P}}X}^{\infty} \frac{\exp\left[-\frac{\nu^2}{2}\right]}{\sqrt{2\pi}} d\nu\right]^K dX \quad . \tag{4-29}$$

It is observed that the distribution $p_1(X_z)$ has an essentially constant spread, but the average value of X_z increases as the square root of n. The spread or variance, $\sigma^2_{X_z}$, goes to zero with the signal power, goes to $1/2S$ as the noise power approaches zero, and never exceeds the signal power. Thus, the variance becomes small, percentagewise, with increasing n. The normalized distribution function p_1' approaches an impulse function.

In view of these properties, one will note that the probability of error given by Eq. (4-8) will not differ much from the result acquired by substituting the average X_z in the expression in brackets, rather than performing the indicated averaging. Then,

$$P(\text{error}) \sim 1 - \left[1 - \int_{\sqrt{\frac{n}{P}}S}^{\infty} p(X)\, dX\right]^K \quad , \tag{4-30}$$

or, by change of variable,

$$P(\text{error}) \sim 1 - \left[1 - \int_{\sqrt{\frac{nS}{P}}}^{\infty} \frac{\exp\left[-\frac{\nu^2}{2}\right]}{\sqrt{2\pi}} d\nu\right]^K \quad . \tag{4-30a}$$

If the integral is small compared to unity (note that if $\sqrt{nS/P}$ is about 1.3, the integral is less than 0.1), a further approximation is given by

$$P(\text{error}) \sim K \int_{\sqrt{\frac{nS}{P}}}^{\infty} \frac{\exp\left[-\frac{\nu^2}{2}\right]}{\sqrt{2\pi}} d\nu \quad . \tag{4-30b}$$

For values of $\sqrt{nS/P}$ consistent with Eq. (4-26), one may use the further approximation

$$\int_Q^{\infty} \exp\left[-t^2/2\right] dt \doteq \frac{1}{Q}\exp\left[-Q^2/2\right]$$

which, when used here, results in the expression

$$P(\text{error}) \sim K\left(\frac{P}{nS}\right)\exp\left[-\frac{nS}{2P}\right] \quad . \tag{4-31}$$

A comparison of Eq. (4-30) with the expression of Eq. (A3-6) shows the expected similarity. Therefore, here too, if the approximations $nS/2P = C'T$ for $S \ll N$, and $K = \exp[H'T]$ for

$K \gg 1$ are used, an expression for the probability of error is

$$P(\text{error}) \sim \frac{1}{2\sqrt{\pi C'T}} \exp\left[-C'T\left(1 - \frac{H}{C}\right)\right] . \qquad (4\text{-}31a)$$

The implication of expressions of the Eq.(4-31a) is that the probability of error, and thus the per-unit equivocation, can be held arbitrarily low by increase of T (involving delay) if only the H/C ratio is less than unity by an amount however small. However, this expression is valid only for small signal-to-noise ratios. For larger values of S/N, the ratio S/P approaches unity and is, of course, smaller than ln(1 + S/N). It follows that H (for large K) must be less than a number smaller than the theoretical capacity if the exponential [ln K – nS/2P] is to remain negative.

It is of incidental interest to note that Shannon shows (see Ref. 22, page 63 – "White Gaussian noise has the peculiar property . . . ") that any random waveform can be used in a system disturbed by white Gaussian noise and can approach the ideal when the signal-to-noise ratios are small. This fact is illustrated in Golay's result[11] for pulse position modulation under the assumption of small signal-to-noise ratio. It is worth repeating that it is the distinct advantage of correlation techniques that they present the most suitable methods of accomplishing the practical identification of randomly varying signals buried in noise.

CHAPTER V

THE THEORETICAL STUDY OF THRESHOLD DETECTION

A. The Probability of Error for an Arbitrary Threshold

Threshold detection is given the distinction of a separate chapter because it represents a technique in common use, and thus deserves an accurate and complete analysis for comparison with the foregoing more-or-less ideal methods of reaching a decision about the transmitted signal.

In review, a threshold-detector scheme is one in which the decision about which of the possible message waveforms was transmitted is based on one of the outputs W_k (see Fig. 3.1) exceeding a predetermined fixed value called the threshold. Of course, if two or more outputs exceed the threshold value, or if none exceeds it, no decision can be made, and an error results. Obviously, there is also an error if only one of the outputs exceeds the threshold but if that output does not correspond to the transmitted message.

A first-order expression of the effect of establishing an arbitrary threshold can be outlined as follows. In Fig. 5.1, for a given received signal $\vec{Z}$, the correlation output W_o is the segment RO times $|\vec{Z}|$. It is assumed that an arbitrary fraction of the expected output W_o is chosen as the threshold, and that OQ is the corresponding component along $\vec{Z}$ necessary to produce the threshold value in the output of a correlator. The expected output is nS, and the threshold is anS, where a is the "threshold coefficient." Thus, the segment OQ corresponds to $aS\sqrt{n/P}$.

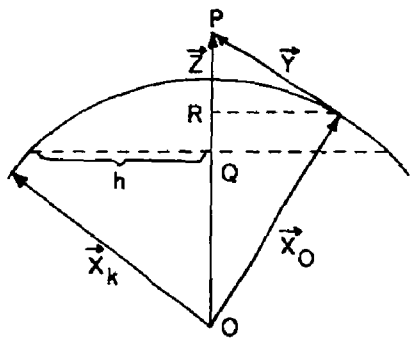

Fig. 5.1. Vector model illustrating threshold detection.

The possibility that a combination of $\vec{X}_0$ and $\vec{Y}$ might occur to yield a Z-component of $\vec{X}_0$ less than OQ is neglected for the moment, and only the probability of error caused by the occasional occurrence of a component of an $\vec{X}_k$ $(k \neq 0)$ in the Z-direction which exceeds OQ will be investigated.

The vectors $\vec{X}_k$ are randomly oriented in n-space, and to a first approximation are of equal length. Thus, they terminate randomly on a hypersphere that is the locus of all points of distance $|X| = \sqrt{nS}$ from the origin. If one of the vectors $\vec{X}_k$ should terminate in the zone above Q in Fig. 5.1, its Z-component will exceed OQ and an error will occur. The probability of errors due to this cause is thus approximately equal to the ratio of the area of the zone to the area of the sphere for any one random vector. The area of the zone will not exceed the area of a hemisphere of radius h (see Fig. 4.1). Thus,

$$P(X_k \text{ lies in zone}) \leq \frac{\frac{1}{2}A(h)}{A(|X|)} \quad . \tag{5-1}$$

Here A(h) is the area of a hypersphere of n dimensions as discussed in Appendix II (see AII-17) and with radius h. From Eq. (5-1), it follows that

$$P(X_k \text{ lies outside zone}) \geq 1 - \frac{h^{n-1}}{2|X|^{n-1}} \quad . \tag{5-2}$$

But h is seen to be $\sqrt{1 - \left[(a^2 nS^2)/P\right]}$, while $|X|$ is the radius of the hypersphere, $\sqrt{nS}$. From these values, Eq. (5-2) may be rewritten

$$P(X_k \text{ lies outside zone}) \geq 1 - \frac{1}{2}\left(1 - \frac{a^2 S}{P}\right)^{\frac{n-1}{2}} \quad , \tag{5-3}$$

or

$$\geq 1 - \frac{1}{2}\left(\frac{N + S(1 - a^2)}{N + S}\right)^{\frac{n-1}{2}} \quad . \tag{5-3a}$$

Since K is always one or greater, one may readily write

$$P(\text{all } X_k \text{ lie outside zone}) \geq \left[1 - \frac{1}{2}\left(1 - a^2 \frac{S}{P}\right)^{\frac{n-1}{2}}\right]K \quad , \tag{5-4}$$

and from Eq. (5-4) the probability of error from this cause (components of wrong vectors exceeding OQ) is

$$P(\text{error}) \leq 1 - \left[1 - \frac{1}{2}\left(1 - a^2 \frac{S}{P}\right)^{\frac{n-1}{2}}\right]K \quad , \tag{5-5}$$

$$< \frac{K+1}{2}\left(1 - a^2 \frac{S}{P}\right)^{\frac{n-1}{2}} \quad , \tag{5-5a}$$

The validity of Eq. (5-5a) is insured by subtracting the smaller value

$$1 - \frac{K}{2}\left(1 - a^2 \frac{S}{P}\right)^{\frac{n-1}{2}}$$

from unity rather than the larger indicated

$$\left[1 - \frac{1}{2}\left(1 - a^2 \frac{S}{P}\right)^{\frac{n-1}{2}}\right]K \quad .$$

Also, K + 1 is certainly larger than the K for which it is substituted. From Eq. (5-5a) it follows that

$$P(\text{error}) < \frac{K+1}{2\sqrt{1 - a^2 \frac{S}{P}}}\left[\left(\frac{N}{N+S}\right)\left(\frac{N + S(1-a^2)}{N}\right)\right]^{\frac{n}{2}} \quad , \tag{5-6}$$

or

$$< \frac{1}{2\sqrt{1 - a^2 \frac{S}{P}}} 2^{\log_2(K+1)\left(\frac{N}{P}\right)^{\frac{n}{2}} + \frac{n}{2}\log_2\left[1 + \frac{S}{N}(1-a^2)\right]} \quad . \tag{5-7}$$

In terms of capacity and information rate of the source, this becomes

$$P(\text{error}) < \frac{1}{2\sqrt{1 - a^2 \frac{S}{P}}} 2^{HT - CT + \frac{n}{2}\log_2\left[1 + \frac{S}{N}(1-a^2)\right]} \quad . \tag{5-8}$$

In Eq. (5-8), increasing the signal duration T will lead to any desired frequency of errors, however small, if CT exceeds $(n/2)\log_2\left[1 + (S/N)(1 - a^2)\right] + HT$. This is equivalent to subtracting from the theoretical capacity a capacity equal to that of a similar system in which the received average power is $(1 - a^2)$ times the original signal power, and requiring that the information rate at the source be less than the diminished capacity. For small signal-to-noise ratios, the remaining useful capacity is only a^2 times the theoretical capacity. From a different point of view, the transmitted power is only a^2 times as effective as the theoretical limit, and a corresponding increase in power might be utilized to regain the performance that would correspond to the ideal use of the original power.

In the foregoing manner, it is shown that an arbitrary choice of a standard threshold will result in a loss of useful channel capacity. However, the careful choice of an optimum threshold will lead to an efficient use of all the system capacity if certain conditions are met, as will be shown subsequently.

To obtain more precise results than the upper limit expressed by Eq. (5-8), it is necessary to include all the conditions under which errors are recorded. A message is correctly indicated when the threshold value is exceeded only by the correct correlation output W_o and not by any other W_k. Any other combination will result in an error.

Therefore, the probability of no error, given for a fixed magnitude of received vector $\vec{Z}$, can be expressed as follows:

$$P(\text{no error}/|Z|) = P(W_o > anS/|Z|)\left[P(W_k < anS/|Z|)\right]^K \quad . \tag{5-9}$$

The probabilities are given as functions of $|Z|$ because it is necessary to compare all W obtained by correlating the various possible signals with a single received signal. The average probability of error is then obtained as an average of the probability of error associated with each received combination of signal and random noise. Thus,

$$P(\text{no error}) = \int_o^\infty p_o(|Z|)\, d|Z|\, P(\text{no error}/|Z|) \quad . \tag{5-10}$$

In Eq. (5-9), the term in brackets is found (in accordance with the methods of Appendix IV) to be given by

$$P(W_k < anS/|Z|) = \int_{-\infty}^{anS} \frac{1}{\sqrt{2\pi S|\vec{Z}|}} \exp\left[-\frac{W_k^2}{2S|Z|^2}\right] dW_k \quad ; \tag{5-11}$$

or

$$P(W_k < anS/|Z|) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{an\sqrt{S}}{|Z|}} \exp\left[-\frac{\nu^2}{2}\right] d\nu \quad . \tag{5-11a}$$

by change of variable.

The probability that W_o exceeds anS is not so readily expressed. However, a large part of the difficulty may be avoided as follows:

$$P(\text{error}) = 1 - \int_o^\infty p_o(|Z|)\, d|Z| \left[1 - P(W_o < anS/|Z|)\right] \left[1 - \int_{\frac{an\sqrt{S}}{|Z|}}^\infty \frac{\exp\left(-\frac{\nu^2}{2}\right)}{\sqrt{2\pi}} d\nu\right]^K \quad , \quad (5\text{-}12)$$

$$\doteq \int_o^\infty p_o(|Z|)\, d|Z| \left[P(W_o < anS/|Z|) + K\int_{\frac{an\sqrt{S}}{|Z|}}^\infty \frac{\exp\left(-\frac{\nu^2}{2}\right)}{\sqrt{2\pi}} d\nu\right] \quad . \quad (5\text{-}12a)$$

In Eq. (5-12a), all terms involving products of the probabilities of individual error are neglected, since in practical cases each is very small, and terms of higher orders of smallness contribute little to the true result. From Eq. (5-12a), one may write

$$P(\text{error}) \doteq P(W_o < anS) + K\int_o^\infty p_o(|Z|)\, d|Z| \int_{\frac{an\sqrt{S}}{|Z|}}^\infty \frac{\exp\left(-\frac{\nu^2}{2}\right)}{\sqrt{2\pi}} d\nu \quad . \quad (5\text{-}13)$$

The first term on the right in Eq. (5-13) is the average probability that W is less than anS and no longer a function of $|Z|$. A study of the probability density distribution, $p_2(W_o)$, which leads to $P(W_o < anS)$, is given in Appendix V. When it is introduced, the probability of error becomes

$$P(\text{error}) \doteq \int_{-\infty}^{anS} p_2(W_o)\, dW_o + K\int_o^\infty p_o(|Z|)\, d|Z| \int_{\frac{an\sqrt{S}}{|Z|}}^\infty \frac{\exp\left(-\frac{\nu^2}{2}\right)}{\sqrt{2\pi}} d\nu \quad . \quad (5\text{-}13a)$$

In Eq. (5-13a), $p_o(|Z|)$ is of the type already frequently used, but stems from a normal distribution with variance $P = N + S$.

The area of $p_2(W_o)$ is distributed about the average value nS, while the standard deviation of the distribution is $\sqrt{nS(N + 2S)}$. Thus, for large values of n and for a somewhat less than unity, the first term on the right of Eq. (5-13a) will be much less than the second (to approximately the extent that K exceeds one). Therefore, the probability of error behaves very much like the second term alone. Furthermore, the final result will not differ greatly from that obtained by substituting the average value of $|Z|$ in the second integral, since this distribution, too, is quite peaked about its average value $\sqrt{nP}$ for large values of n. It follows that for n large and for a in the range $0 < a < 0.8$, one may write

$$P(\text{error}) \sim K\int_{\frac{an\sqrt{S}}{\sqrt{nP}}}^\infty \frac{\exp\left(-\frac{\nu^2}{2}\right)}{\sqrt{2\pi}} d\nu \quad . \quad (5\text{-}14)$$

When the approximation used in Eq. (4-30) is applied here, there results

$$P(\text{error}) \sim \frac{K}{\sqrt{2\pi a}} \sqrt{\frac{P}{nS}} \exp\left(-a^2 \frac{nS}{2P}\right) \quad . \quad (5\text{-}15)$$

To gain still further insight into the meaning of this expression (but subject to the conditions of K >> 1 and S << N), other substitutions as used in Chapter IV lead to the result

$$P(\text{error}) \sim \frac{1}{2\sqrt{\pi a^2 C'T}} \exp\left[-a^2 C'T\left(1 - \frac{H}{a^2 C}\right)\right] \quad . \tag{5-16}$$

The expression, subject to the same conditions discussed under the inequality (5-8), affirms the fact that the probability of error may be made arbitrarily small with increasing T if the ratio of information generated at the source to the theoretical system capacity does not exceed the square of the relative threshold coefficient a. This decrease of effective capacity is made apparent by the appearance of a^2C in each place that C appears in Eq. (4-31).

B. The Optimum Threshold

As developed in the preceding section, the probability of error when a threshold criterion of detection is used can be written

$$P(\text{error}) = 1 - \int_0^\infty p_0(|Z|)\, d|Z| \left[\int_{anS}^\infty p_3(W_o/|Z|)\, dW_o\right]\left[\int_{-\infty}^{anS} p_4(W_k/|Z|)\, dW_k\right]^K \quad . \tag{5-17}$$

The density functions $p_3(W_o/|Z|)$ and $p_4(W_k/|Z|)$ have not been used previously in this paper, and they represent merely the results of differentiating with respect to the appropriate parameter of the probabilities $P(W_o > anS/|Z|)$ and $P(W_k < anS/|Z|)$, respectively. To determine the optimum value of threshold coefficient a which leads to the minimum probability of error, P(error) is differentiated with respect to a, and the resulting function examined for zeros.

$$\frac{\partial P(\text{error})}{\partial a} = -\int_0^\infty p_0(|Z|)\, d|Z| \left\{\left[\int_{-\infty}^{anS} p_4(W_k/|Z|)\, dW_k\right]^K \left[-p_3(anS)\right]\right.$$
$$\left. + \left[\int_{anS}^\infty p_3(W_o/|Z|)\, dW_o\right] K \left[\int_{-\infty}^{anS} p_4(W_k/|Z|)\, dW_k\right]^{K-1} p_4(anS)\right\} \quad , \tag{5-18}$$

which is zero if

$$p_3(anS)\int_{-\infty}^{anS} p_4(W_k/|Z|)\, dW_k = p_4(anS)\, K \int_{anS}^\infty p_3(W_o/|Z|)\, dW_o \quad . \tag{5-19}$$

Thus, a zero of Eq. (5-18) and a minimum of P(error) occurs when

$$p_3(anS) = p_4(anS) \cdot K \frac{\int_{anS}^\infty p_3(W_o/|Z|)\, dW_o}{\int_{-\infty}^{anS} p_4(W_k/|Z|)\, dW_k} \quad . \tag{5-20}$$

Since $p_4(W_k/|Z|)$ is a modified normal distribution about zero mean, and $p_3(W_o/|Z|)$ is distributed about a mean value nS, for all a between zero and unity, the ratio of integrals in Eq. (5-20) must lie between 1/2 and 2. When n is large, the peaking of the distributions leads to a ratio approaching unity. Thus, for large n, one may write

$$p_3(anS) \doteq K p_4(anS) \quad . \tag{5-21}$$

One may make use of the fact that the variance of variable W_o is given by $nS(N + 2S)$, while that of W_k is nSN. To a first-order approximation, if the p's are represented by normal distributions, Eq. (5-21) becomes

$$\frac{1}{\sqrt{2\pi nS(N + 2S)}} \exp\left[-\frac{(anS - nS)^2}{2nS(N + 2S)}\right] = K \frac{1}{\sqrt{2\pi nSN}} \exp\left[-\frac{(anS)^2}{2nSN}\right] , \tag{5-22}$$

or

$$\frac{1}{\sqrt{N + 2S}} \exp\left[-\frac{(1 - a)^2 nS}{2(N + 2S)}\right] \doteq K \frac{1}{\sqrt{N}} \exp\left[-\frac{a^2 nS}{2N}\right] . \tag{5-22a}$$

From Eq. (5-22a), it is evident that

$$\frac{(1 - a)^2}{N + 2S} = \frac{a^2}{N} - \frac{2}{nS} \ln K \sqrt{\frac{N + 2S}{N}} . \tag{5-22b}$$

The solutions of this equation are

$$a = \frac{N}{2S} \pm \sqrt{\left(\frac{N}{2S}\right)^2 + \frac{N}{2S}\left[1 + \frac{2}{n}\left(2 + \frac{N}{S}\right)\ln K + \frac{1}{n}\left(2 + \frac{N}{S}\right)\ln\left(1 + \frac{2S}{N}\right)\right]} . \tag{5-23}$$

When n is very large (which, incidentally, is necessary if the normal approximation to the functions p_3 and p_4 is to be valid) and for $S/N \ll 1$, the solution becomes

$$a \doteq \frac{N}{2S}\left[\left(1 + \frac{S}{N} + \cdots\right)\left(1 + \frac{2}{n} \ln K + \cdots\right) - 1\right] , \tag{5-24}$$

or

$$a \doteq \frac{N}{2S}\left[\frac{S}{N} + \frac{2}{n} \ln K\right] . \tag{5-24a}$$

When the indicated operations are carried out,

$$a \doteq \frac{1}{2}\left[1 + \frac{2}{np} \ln K\right] . \tag{5-24b}$$

Under the conditions of $K \gg 1$, and those outlined before, such that the approximations $1/2\,np \doteq C'T$ and $\ln K \doteq H'T$, Eq. (5-24b) may be written

$$a \doteq \frac{1}{2}\left[1 + \frac{H}{C}\right] . \tag{5-25}$$

This expression has appeared in slightly different form in the note by Golay,[11] and is the result obtained by all such optimizations of small signals detected by the threshold criterion of detection. The important point brought out by this result is that if one is to obtain the maximum use of a given channel when threshold detection is employed (subject to the conditions listed above for the validity of the result), he must adjust the threshold coefficient quite close to unity. In fact, the difference between H/C and unity must be approximately halved by the difference between the threshold coefficient and unity for the optimum result. Equation (5-25) also illustrates the fact that, if a channel is to be used to carry information at a rate substantially less than the theoretical channel capacity, there exists a definite threshold for which the errors are a minimum.

When the optimum value of a is placed in Eq. (5-16), the result is

$$P(\text{error}) \sim \frac{1}{\sqrt{\pi C'T\left(1+\frac{H}{C}\right)}} \exp\left[-\frac{C'T}{4}\left(1-\frac{H}{C}\right)^2\right] \quad . \tag{5-26}$$

It is emphasized that expression (5-26) is not valid for very near unity (and thus H/C near unity) but is significant, however, for H/C of the order of one-half or so.

A further study of the effects introduced by the use of a standard threshold for correlation detection can be made by postulating a highly idealized system of the following type.*

A set of K + 1 vectors is arranged in n-space so that they are mutually orthogonal. Obviously, n must exceed K. The radius of the hypersphere on which all K + 1 vectors terminate will be $|X| = \sqrt{nS}$. Under these conditions, a random orientation of the vector system in a set of Cartesian coordinates of the n-space will cause the coordinate values to be measured by numbers which very nearly (but not exactly) follow a Gaussian distribution with a variance S. The coordinates are used to generate the K + 1 possible message waveforms.

As before, one of these vectors (called $\vec{X}_0$) is selected and transmitted, noise being added in the channel through which transmission occurs. At the receiver, the received signal $\vec{X}_0 + \vec{Y}$ is cross-correlated with each of the K + 1 vectors $\vec{X}_k$, and the outputs are used in arriving at a decision about which of the messages was transmitted. An error will occur in all cases in which the noise component added to $\vec{X}_0$ is less than (a – 1) $\sqrt{nS}$ (in which case $|X_o|^2 + \vec{X}_0 \cdot \vec{Y} < anS$), or in which the component along any of the K other vectors exceeds a $\sqrt{nS}$ (in which case $\vec{X}_k \cdot \vec{X}_0 + \vec{X}_k \cdot \vec{Y}$ equals $\vec{X}_k \cdot \vec{Y} > anS$), or any combination of these events. By the method of Appendix IV, if the noise vector $\vec{Y}$ arises from a Gaussian distribution, we may take the components along any of the coordinate axes, or along any of the mutually orthogonal vectors, as numbers from the same Gaussian distribution. Therefore,

$$P(\text{no error}) = \int_{(a-1)|X|}^{\infty} p(Y)\,dY \left[\int_{-\infty}^{a|X|} p(Y)\,dY\right]^K \quad , \tag{5-27}$$

$$= \left[1 - \int_{-\infty}^{(a-1)|X|} p(Y)\,dY\right]\left[1 - \int_{a|X|}^{\infty} p(Y)\,dY\right]^K \quad . \tag{5-27a}$$

The probability is the product of the probabilities of individual events because the coordinates of the noise vector are independent.

Since

$$\int_{a|X|}^{\infty} p(Y)\,dY$$

should be quite small in practical cases, it follows that

$$P(\text{error}) \doteq 1 - \left[1 - \int_{(1-a)|X|}^{\infty} p(Y)\,dY\right]\left[1 - K\int_{a|X|}^{\infty} p(Y)\,dY\right] \quad . \tag{5-28}$$

Again neglecting terms of the second order of smallness, Eq. (5-28) becomes

*This system and discussion is substantially the same as that appearing in a PROJECT LINCOLN internal memorandum by the author, "Cross Correlation Thresholds and Channel Capacity," 10 September 1951.

$$P(\text{error}) \doteq \int_{(1-a)|X|}^{\infty} p(Y)\,dY + K\int_{a|X|}^{\infty} p(Y)\,dY \quad . \tag{5-29}$$

When the optimum value for a is substituted (and large K and n are assumed), the form of Eq. (5-29) is

$$P(\text{error}) \doteq \int_{\frac{1}{2}\left(1-\frac{H}{C}\right)\sqrt{n\rho}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\nu^2}{2}\right) d\nu + \int_{\frac{1}{2}\left(1+\frac{H}{C}\right)\sqrt{n\rho}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\nu^2}{2}\right) d\nu \quad , \tag{5-30}$$

The approximation

$$\int_{Q}^{\infty} \exp\left(-\frac{1}{2}t^2\right) dt = \frac{1}{Q}\exp\left(-\frac{Q^2}{2}\right) \quad ,$$

valid for large Q, can be used here when $n\rho$ is large. Then,

$$P(\text{error}) \doteq \frac{1}{\sqrt{2\pi}}\left[\frac{2}{\left(1-\frac{H}{C}\right)\sqrt{n\rho}}\right]\exp\left[-\frac{1}{8}\left(1-\frac{H}{C}\right)^2 n\rho\right]$$
$$+\frac{K}{\sqrt{2\pi}}\left[\frac{2}{\left(1+\frac{H}{C}\right)\sqrt{n\rho}}\right]\exp\left[-\frac{1}{8}\left(1+\frac{H}{C}\right)^2 n\rho\right] \quad . \tag{5-31}$$

But K has already been assumed large, and if $S \ll N$ so that $1/2\, n\rho = C'T$, the expression becomes

$$P(\text{error}) \doteq \frac{\exp\left[-\frac{1}{4}C'T\left(1-\frac{H}{C}\right)^2\right]}{\sqrt{\pi C'T}\left(1-\frac{H}{C}\right)} + \frac{\exp\left[-\frac{1}{4}C'T\left(1+\frac{H}{C}\right)^2 + H'T\right]}{\sqrt{\pi C'T}\left(1+\frac{H}{C}\right)} \quad , \tag{5-32}$$

or

$$P(\text{error}) \doteq \frac{2}{\sqrt{\pi C'T}\left[1-\left(\frac{H}{C}\right)^2\right]}\exp\left[-\frac{C'T}{4}\left(1-\frac{H}{C}\right)^2\right] \quad . \tag{5-33}$$

The expressions given by (5-31) and (5-33) are valid only when the threshold coefficient is optimized (and subject to the conditions listed). A comparison of these with (4-30) and (4-31) reveals two differences of primary consideration. One of these differences, the fact that the ratio S/P appears in Chapter IV while S/N appears in the results just obtained, is easily explained by noting that P = N + S includes the variation of the magnitudes of the signal vectors. These vectors were of fixed length in the example just given. The other difference is the presence of the squared factor $(1 - H/C)$ in the case of threshold detection while only the first power of that factor appears in the case of maximum correlation detection. This means that, for the same conditions of noise, capacity and information rate, considerably greater delay is required in the case of threshold detection to obtain a given probability of error than when the maximum correlation criterion is used.

CHAPTER VI

NOISE IN THE AUXILIARY CHANNELS

Up to this point, the systems considered have been those in which some auxiliary noiseless channels have been utilized to make the set of possible message waveforms available at the receiver for cross-correlation with the received signal. Such systems are, of course, special cases of a more general class of systems in which the auxiliary channels are also disturbed by unavoidable additive noise.

The prime effect of such disturbing noise in the auxiliary channels is to introduce some doubt as to precisely what it is that is being sought in the cross-correlation process at the receiver. An additional uncertainty remains after a message has been received and the correlation outputs obtained over that present when the auxiliary channels are noiseless. Because of this greater uncertainty, it is expected that the decisions made on the basis of these correlator outputs will be more frequently in error than when the noise is absent. This is found to be the case.

As in the earlier systems discussed, it is here considered that the message $\vec{X}_0$ has been transmitted through the intelligence channel in which noise $\vec{Y}$ is added. Also, the set of possible messages $\vec{X}_k$ is transmitted in auxiliary channels in which noises $\vec{Y}_k$ are added. It is assumed that the noises $\vec{Y}_k$ $(k \neq 0)$ are independent of noise $\vec{Y}_0$.

At the receiver, $\vec{Z}$, the received signal, is cross-correlated with each of the set of $\vec{Z}_k$ waveforms representing the corrupted versions of the possible message waveforms. The outputs resulting from the correlations are made the basis of a decision about which of the message waveforms was transmitted.

At the receiver, the correlation of $\vec{Z}$ with $\vec{Z}_k$ will yield

$$\vec{Z} \cdot \vec{Z}_k = (\vec{X}_0 + \vec{Y}) \cdot (\vec{X}_k + \vec{Y}_k) \quad , \tag{6-1}$$

$$= \vec{X}_0 \cdot \vec{X}_k + \vec{Y} \cdot (\vec{X}_k + \vec{Y}_k) + \vec{X}_0 \cdot \vec{Y}_k \quad . \tag{6-1a}$$

If the signal-to-noise ratio is ρ in the intelligence channel and S/N_k in the auxiliary channels, there will again be an expected value nS in the correct correlator output, while zero is expected in all the other outputs. Geometrically, as illustrated in Fig. 6.1, the components of $\vec{Z}_k$ that fall along $\vec{Z}$ will follow a Gaussian distribution with variance $S + N_k = P_k$, except when $k = 0$. For simplicity, all N_k are considered equal in the discussion following.

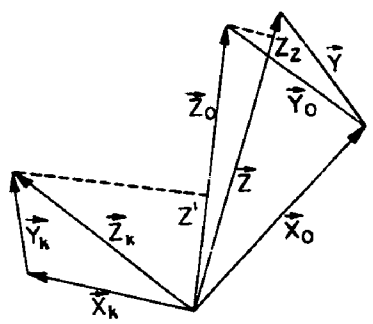

Fig. 6.1. Vector model illustrating noise in the auxiliary channel.

Let $p_5(Z_z)$ be the distribution of the components of $\vec{Z}_0$ along $\vec{Z}$. Note that Z_z corresponds to X_z of the systems discussed in Chapters IV and V.

When the criterion of detection is that of maximum correlation, the probability of error

takes the form

$$P(\text{error}) = 1 - \int_{-\infty}^{\infty} p_5(Z_z)\, dZ_z \left[\int_{-\infty}^{Z_z} p(Z')\, dZ' \right]^K \quad , \tag{6-2}$$

where Z' is the component of $\vec{Z}_k$ along $\vec{Z}$. Of course, Z_z is the component of $\vec{Z}_0$ along $\vec{Z}$.

Maximum correlation is used as an example for comparison with earlier results. It is evident that the contribution to the outputs from the correlators that do not correspond to the transmitted waveform will be a function of the signal plus noise power P_k rather than the signal power S alone. This is true independently of whether the criterion of detection is that of maximum correlation or exceeding a threshold. Consequently, the final result obtained for noise added in the auxiliary channel can be extended readily to threshold detection by analogy.

The distribution $p_5(Z_z)$ is peaked about its average value in much the same manner as $p_1(X_z)$. This average value may be obtained as follows. The expected output is NS. The average magnitude of $\vec{Z}_0$ is $\sqrt{nP}$. Thus, although $\vec{Z}_0$ and Z_z are not independent, the distributions of these quantities for large n are sufficiently peaked that, to a good approximation,

$$\text{ave } Z_z = \frac{nS}{\sqrt{nP}} = S\sqrt{\frac{n}{P}} \quad . \tag{6-3}$$

The behavior of the probability of error, again under the assumption of large n, may be obtained by evaluating Eq. (6-2) at the average value of Z_z. Therefore,

$$P(\text{error}) \sim K \int_{S\sqrt{\frac{n}{P}}}^{\infty} \frac{\exp\left[-\frac{\nu^2}{2P_k}\right]}{\sqrt{2\pi P_k}}\, d\nu \quad , \tag{6-4}$$

$$\sim \frac{K}{\sqrt{2\pi}} \sqrt{\frac{PP_k}{nS^2}} \exp\left[-\frac{nS^2}{2PP_k}\right] \quad . \tag{6-4a}$$

Again use is made of the approximations for $\rho << 1$ and $K >> 1$ which are, respectively, $1/2\, n\rho \doteq C'T$ and $\ln K = H'T$. Equation (6-4a) becomes

$$P(\text{error}) \sim \frac{1}{2\sqrt{\pi C'T}} \sqrt{\frac{P_k}{S}} \exp\left[-C'T\left(\frac{S}{P_k} - \frac{H}{C}\right)\right] \quad , \tag{6-5}$$

$$\sim \frac{1}{2\sqrt{\pi \frac{S}{P_k} C'T}} \exp\left[-\frac{S}{P_k} C'T\left(1 - \frac{H}{\frac{S}{P_k} C}\right)\right] \quad . \tag{6-5a}$$

Note that for $N_k = 0$, $S/P_k = 1$, and Eq. (6-5) reduces to the result obtained in Chapter IV see Eq. (4-31). Furthermore, to obtain arbitrarily small probability of error,

with increasing T, it is indicated that the ratio H/C should be less than the signal-to-noise ratio in the auxiliary channels. The appearance of $(S/P_k)C'$ in each place where C' appears in Eq. (4-31) indicates that the effective system capacity is S/P_k times the theoretical capacity. It may be expressed by stating that communication is possible, relatively speaking, to the extent that noise is absent in the auxiliary channels.*

*It is suggested that transmission of the "code book," which is somewhat akin to learning, must be accomplished perfectly if the information presented in that code is to be transmitted at a rate equal to the capacity of the channel. Also, since it is evident that some form of auxiliary channel must be used for what amounts to code-book transmission, the total equivalent capacity can never be utilized perfectly.

CHAPTER VII

THE EXPERIMENTAL STUDY OF PROBABILITY OF ERROR

A. The Modified Theory for K = 1

In the experimental verification of the theoretical material presented in this paper, a relatively simple physical system was constructed. In this system, one possible message waveform is generated in the transmitter and sent through an auxiliary channel to the receiver. The waveform is also sent through the intelligence channel, modulated in an on-off manner, and, in the process of transmission, disturbing noise is added. A simplified block diagram of the system is shown in Fig. 7.1, and a detailed description of the actual system follows in a later section of this chapter. The system allows for the possible introduction of a disturbing noise in the auxiliary channel also.

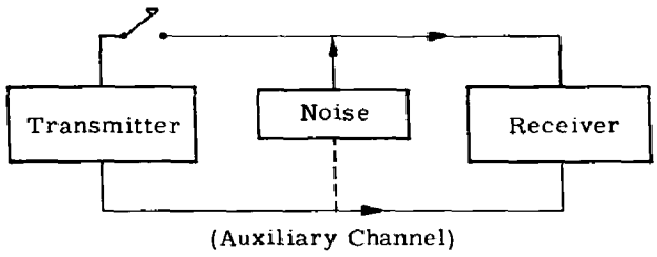

Fig. 7.1. Simplified block diagram of the binary-choice system.

It is understood that in this system K + 1 = 2, or K = 1, the alternative signal to the waveform generated at the transmitter being a zero or null waveform corresponding to the key-up position. In the experimental system, the on-off indicator at the receiver operates to indicate "on" whenever a preset threshold is exceeded by the correlation receiver.

Certain simple modifications of the analysis underlying the general case for large K are required to theoretically describe this NOMAC system. Here, an error is made when either the correlation for the key-up position exceeds the threshold or when, for the key-down position, the threshold is not exceeded.

Let the threshold be set at X^*. On the average, the probability of error is given by

$$P(\text{error}) = P(\text{on})\left[P\left(|X|^2 + |X|\,Y_1 < X^*\right)\right] + p(\text{off})\left[P\left(|X|\,Y_1 > X^*\right)\right] \quad . \tag{7-1}$$

This equation can be expressed as the average over all $|X|$ of the probability of error when $|X|$ is fixed. Thus,

$$P(\text{error}) = \int_0^\infty p_0(|X|)\, d\,|X| \left[P(\text{on}) \int_{|X|^2 - X^*}^\infty \frac{\exp\left[-\frac{\nu^2}{2N\,|X|^2}\right]}{\sqrt{2\pi N}\,|X|} d\nu + P(\text{off}) \int_{X^*}^\infty \frac{\exp\left[-\frac{\nu^2}{2N\,|X|^2}\right]}{\sqrt{2\pi N}\,|X|} d\nu \right]. \tag{7-2}$$

If the system is being used efficiently, the symbols "on" and "off" will occur with equal probability. Then,

$$P(\text{error}) = \frac{1}{2}\int_0^\infty p_0(|X'|)\, d\,|X| \left\{ \int_{|X|^2 - X^*}^\infty \frac{\exp\left[-\frac{\nu^2}{2N\,|X|^2}\right]}{\sqrt{2\pi N}\,|X|} d\nu + \int_{X^*}^\infty \frac{\exp\left[-\frac{\nu^2}{2N\,|X|^2}\right]}{\sqrt{2\pi N}\,|X|} d\nu \right\} . \quad (7\text{-}2a)$$

A choice of X^*, the threshold, is made to minimize the probability of error. It follows that the solution of

$$\frac{\partial}{\partial X^*} P(\text{error}) = 0 \quad (7\text{-}3)$$

is taken for the threshold. A solution of Eq.(7-3) is a solution of

$$-\frac{1}{\sqrt{2\pi N}\,|X|} \exp\left[-\frac{(|X|^2 - X^*)^2}{2N\,|X|^2}\right] = -\frac{1}{\sqrt{2\pi N}\,|X|} \exp\left[-\frac{(X^*)^2}{2N\,|X|^2}\right] , \quad (7\text{-}4)$$

or

$$|X|^2 - X^* = X^*, \quad X^* = \frac{|X|^2}{2} , \quad (7\text{-}5)$$

where $|X|$ is fixed.

The threshold value as considered here is a constant and not a function of $|X|$. To obtain that constant, X^* is averaged over all $|X|$ and the result nS/2 is obtained. This is the constant value of optimum threshold used. Then,

$$P(\text{error}) = \int_0^\infty p_0(|X|)\, d\,|X| \int_{\frac{nS}{2}}^\infty \frac{\exp\left[-\frac{\nu^2}{2N|X|^2}\right]}{\sqrt{2\pi N}\,|X|^2} d\nu , \quad (7\text{-}6)$$

or

$$P(\text{error}) = \int_0^\infty p_0(|X|)\left\{\frac{1}{2}\,\text{erf}\left(\frac{nS}{2\sqrt{N}\,|X|}\right)\right\} d\,|X| . \quad (7\text{-}6a)$$

The probability distribution function $p_0(|X|)$ is the chi-squared type discussed in detail in Appendix II.

If the substitution $|X| = a\sqrt{nS}$ is made, Eq.(7-6) reduces to a more convenient form (in which Stirling's approximation for $\Gamma(n)$ has been used), namely,

$$P(\text{error}) = \sqrt{\frac{n}{\pi}} \int_0^\infty a^{n-1} \exp\left[-\frac{n}{2}(a^2 - 1)\right]\left\{\frac{1}{2}\,\text{erf}\sqrt{\frac{n\rho}{4a}}\right\} da . \quad (7\text{-}7)$$

Erf(Q) is defined here as

$$\sqrt{\frac{2}{\pi}} \int_Q^\infty \exp\left[-\frac{t^2}{2}\right] dt .$$

Another form of Eq.(7-7), using the nomenclature of Fano, reads

$$P(\text{error}) = \sqrt{\frac{n}{\pi}} \int_0^\infty a^{n-1} \exp\left[-\frac{n}{2}(a^2 - 1)\right]\left\{\frac{1}{2}\,\text{erf}\sqrt{\frac{E_s}{2aN_0}}\right\} da . \quad (7\text{-}7a)$$

The argument of the error function is then the square root of the modified ratio of the signal energy to the noise power per cycle.

Curves of the probability of error for several typical values of time-bandwidth product and signal-to-noise ratio are given in Fig. 7.2. These results were obtained by numerical integration of Eq.(7-7), and the method as well as tabulated results are shown in Appendix VI. The probability of error is seen to be primarily a function of the product $n\rho$, and is plotted as a function of that parameter in Fig. 7.3. Over the range of values of $n\rho$ for which the plot is given in Fig. 7.3, the deviation from the single line is less than the line width of the curve for $n = 1000$ when $n = 100$ and more. Where its values differ from the other curve by an appreciable amount, the separate curve for $n = 10$ is shown.

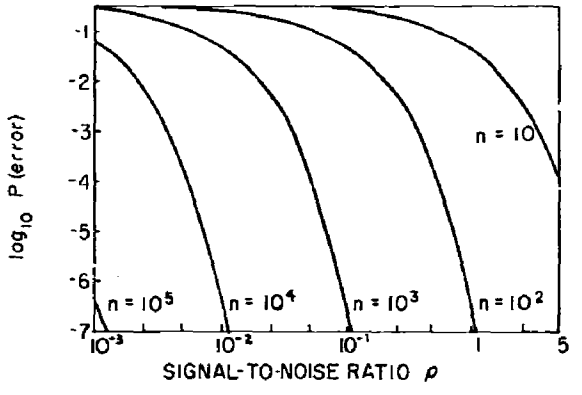

Fig. 7.2. Probability of error as a function of signal-to-noise ratio (ρ) and system bandwidth ratio (n).

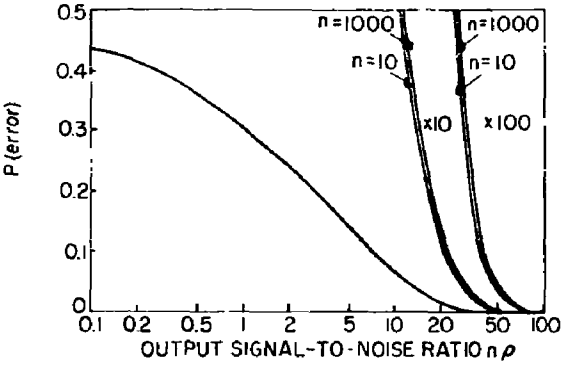

Fig. 7.3. Probability of error as a function of $n\rho$.

When it is recalled that $n\rho$ represents the signal-to-noise ratio at the output of a correlation detector (as well as the ratio E_s/N_o), the significance of the fact that the probability of error is primarily a function of this output ratio becomes apparent. It justifies the assumption that the noise component of the output signal is Gaussian, because the curve as a function of $n\rho$ is essentially that approached by Eq.(7-7a) for n large, which is in turn the error function evaluated at $\alpha = 1$. The error function is, of course, expressed in terms of the Gaussian or normal distribution.

B. Description of the Equipment

In its physical form, the equipment required to investigate the validity of the theoretical results consists of five panels mounted, together with power supplies, on a single relay rack. In addition, the Davenport probability-measuring equipment[3] was used to make actual measurements of the results experimentally. The complete experimental setup is pictured in Fig. 7.4.

On the relay rack containing the NOMAC system (at left in the picture), the following are mounted (from top to bottom): low-voltage power supply, random (noise-like) signal source, side-band generator (transmitter), the simulated channel, the correlation converter, the

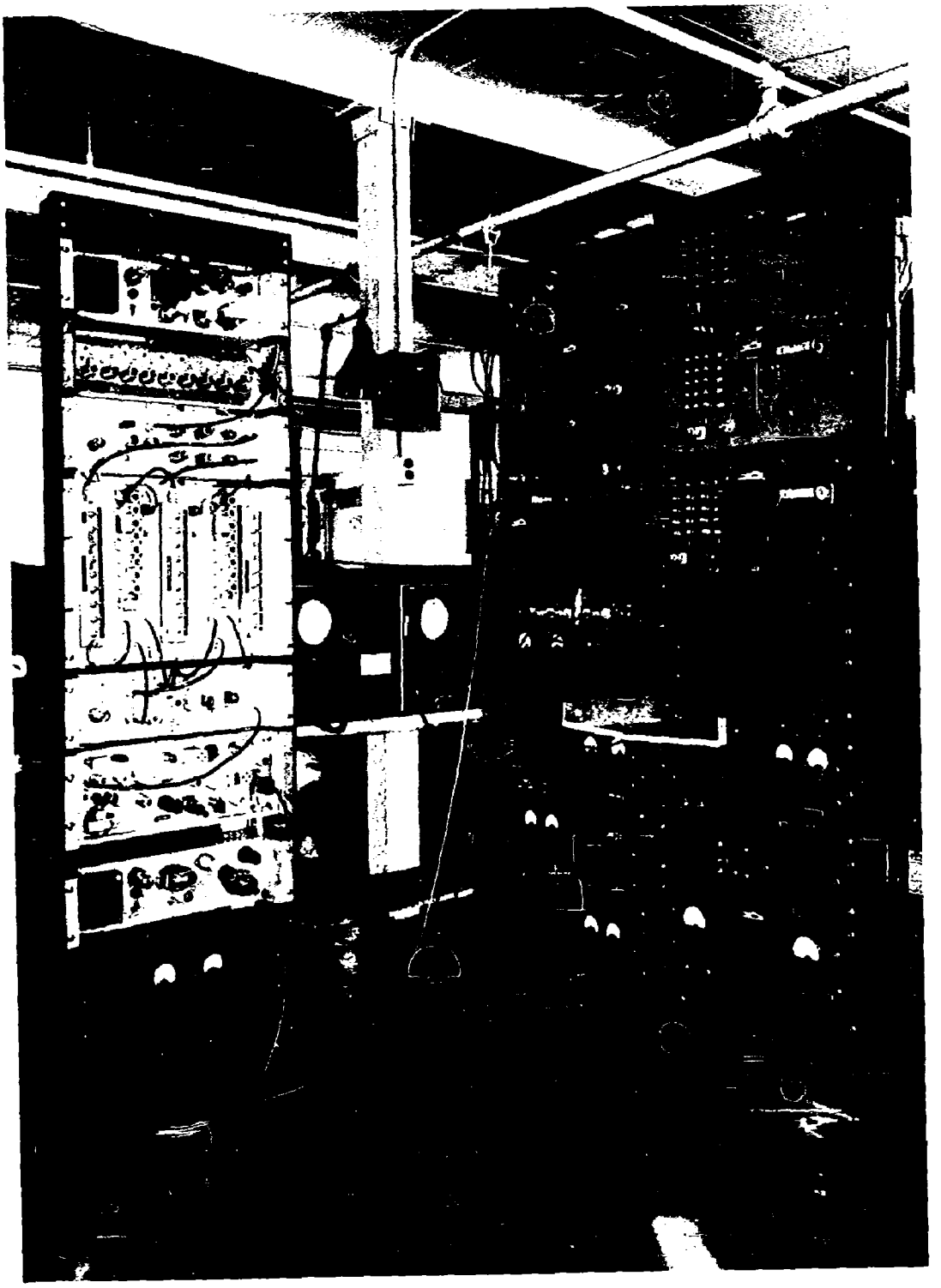

Fig. 7.4. Photograph of the experimental equipment.

integrator-detector (receiver), the channel-noise power supply, and the receiver-transmitter power supply. The interconnection of components other than power supplies is shown in the block diagram (Fig. 7.5).

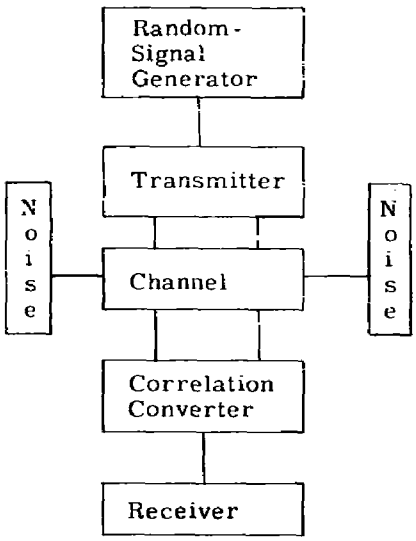

Fig. 7.5. Block diagram of the experimental system.

The schematic diagram of the random-signal source is given in Fig. A7.2 of Appendix VII. Essentially, it is a filter amplifier with a bandwidth of 5 megacycles and a midband frequency of 35 megacycles. The amplifier, featuring an 8-pole Tchebychef frequency response, was designed and built by R. Price and W. McLaughlin of Group 34 of PROJECT LINCOLN. The gain of the 8 stages (6BH6's) is sufficient to amplify the thermal noise and tube noise of the first stage to an output power of about 3 milliwatts of white noise into a 75-ohm load impedance. The noise output 3 Mc from the midband frequency is down 50 decibels or more from the level within the pass band.

The side-band generator or transmitter is shown schematically in Fig. A7.3. The signal input from the random-signal source is mixed with a 5.35-Mc signal from the crystal-controlled local oscillator. Separate 2-stage band-pass amplifiers select the resulting side bands at 29.65 Mc and 40.35 Mc. Each of these stagger-tuned amplifiers is adjusted to an effective bandwidth of 2.0 Mc, and fed through cathode followers to the simulated channel. The resulting side-band spectra have rounded corners so that the transition from channel noise to channel-plus-signal noise will be a smooth function of frequency, and thus less apparent to an unfriendly receiver.

The simulated channel consists of attenuators for the two signals and for the noise added in as the corrupting or disturbing noise. Also, an adding circuit is used to obtain the sum of the proper amounts of random signal and contamination noise. In the first part of the experimental work, the 40.35-Mc signal is connected directly to the correlator (as the reference signal) to simulate a storage process which makes a noise-free version of the signal available for cross-correlation. In the second part, noise is added in the 40.35-Mc channel also, to simulate the conditions when both reference signal and intelligence signal must be transmitted through a noisy medium. The schematic diagram is shown in Fig. A7.4. The two sources of disturbing noise are modified radar IF strips, originally built for the M.I.T. Radiation Laboratory. When operated at full gain and with no input, these strips have a noise output of approximately one volt.

The correlation converter comprises a pair of wide-band RF amplification stages, one each at 29.65 Mc and 40.35 Mc, and a 6AS6 used as a multiplier. The choice of the 6AS6 was made because the same bias on grids one and three, namely, −2.5 volts, places the tube in an operating range leading to a minimum of distortion in the multiplication process. Thus the biasing problem was simplified. The output of this tube is fed to a two-stage 10.7 Mc IF amplifier using conventional components.

The frequency of the IF amplifier is chosen as follows: when the signals at the grids of the 6AS6 are $n_1(t)\cos[\omega_1 t + \theta_1(t)]$ and $n_2(t)\cos[\omega_2 t + \theta_2(t)]$, the product $\psi_{12}(t)$ is given by

$$\psi_{12}(t) = n_1(t)\,n_2(t)\,\frac{1}{2}\left\{\cos[(\omega_2 - \omega_1)t + \theta_2(t) - \theta_1(t)] + \cos[(\omega_2 + \omega_1)t + \theta_2(t) + \theta_1(t)]\right\} \quad . \tag{7-8}$$

Now if n_1 and n_2 as well as the phases θ_1 and θ_2 are independent, ($n_1(t)$ and $n_2(t)$ have zero average), the average of $n_1 n_2$ is zero and the phase of the term at frequency $\omega_1 - \omega_2$ is random. However, if $n_1 = n_2$ and $\theta_1 = \theta_2$, there is no random phase in the difference-frequency term and the product $n_1(t)^2$ has an average proportional to the power (or variance). Thus, in the experimental model, an output of the multiplier at the difference frequency of 10.7 Mc should occur whenever the two side bands generated in the transmitting process are applied to the two signal grids. Figure A7.5 is the schematic diagram of the correlation converter.

The schematic diagram of Fig. A7.6 shows that the integrator-detector is merely a double-conversion superheterodyne receiver using a crystal-controlled oscillator at the first converter and a Collins Type 70E-15 oscillator for the second-converter oscillator. The IF frequency is 455 kc and, with a crystal filter at the same frequency, integration bandwidths of the order of 80 cycles are obtained. In the broad-band (crystal out) position, the bandwidth is 4000 cycles. The former bandwidth is typical of the requirements for teletype circuits, wnile the latter is adequate for speech communication. Filters, as integration devices, are discussed in Chapter VIII.

C. Discussion of Experimental Results

An examination of Eqs.(7-2a) and (7-6) reveals that, for the conditions imposed, the probability of error for "on-off" signaling is the same as the relative frequency of errors in receiving only "on" signals, since the individual integrals contribute equally to the probability of error. This somewhat simplifies the experimental procedure in that measurements or counts made with the signal present in the intelligence channel can be taken as the desired probability of error.

The probability-distribution analyzer developed by W. B. Davenport, Jr. was used to count the relative number of times the output signal exceeded the threshold value which was set at one-half the average output value. The difference between this relative count and unity was taken as the experimental probability of error.

Curves were taken with the signal-to-noise ratio at the input as the independent variable, and various values of the signal-to-noise ratio in the auxiliary channel as a parameter for the two values of bandwidth ratio. When the auxiliary-channel signal-to-noise ratio is infinite, the measured values of probability of error should check the theoretical probability of error as obtained by numerical integration and presented in Fig. 7.2. When noise is added in the auxiliary channel, the effect as predicted in Chapter VI should be observed. This effect is to change the effective value of $n\rho$ in the intelligence channel by the factor $S/(N_k + S)$. This means the theoretical probability-of-error curves should be shifted by an amount corresponding to that factor.

The curves in Figs. 7.6, 7.7, and 7.8 show the results of the experimental measurements. Figure 7.6 shows the probability-of-error measurements made with the wide-band filter.

The value of n calculated for this filter is 615, based on the equivalent noise bandwidths of the input (RF) and output (IF) filters. The measurements were made for probability of error less than 10 per cent which corresponds to an output (postcorrelation) signal-to-noise ratio of about +8 db. This upper limit was established more or less arbitrarily because of the type detection used, i.e., envelope detection, after multiplication and filtering. The level of 8 db represents a signal-peak-voltage to rms-noise-voltage ratio of about 2.5, for which the number of excursions of noise voltage which exceed the peak signal voltage is of the order of one per cent if the noise is approximately Gaussian, and the portion of time spent by the combined signal and noise voltage in the curved portion of the detector characteristic is correspondingly small. Figure 7.7 shows similar results for the narrow-band filter which has a measured n of 31,700. In practice, this filter is so narrow that the drift of the crystal-controlled oscillators is a serious problem, and readings are difficult since all tuning and level settings must be corrected before a count and then checked for drift after the count. For this reason, only one additional curve for noise in the auxiliary channel was taken to show again the predicted displacement. It is to be assumed that the effect shown in Fig. 7.6 would have appeared here also for other values of auxiliary-channel signal-to-noise ratio.

Figure 7.8 shows the results of the earlier figure, plus some additional counts, as a function of "effective nρ," where the effective value of nρ is defined as $n(S/P)_{aux}$. The composite theoretical curve is given for comparison of theoretical with experimental results.

The spread of the points in Fig. 7.8 is attributed to two principal sources of error. The ability to match signal and noise powers was limited by the resolution of the attenuators, which had a range from zero to 101 decibels attenuation in steps of one decibel. The accuracy of components used in the attenuators was 5 per cent, but in general the accuracy of any given step was considerably better than that. However, the absolute error introduced by the larger steps was observed to be as much as 0.4 decibels.

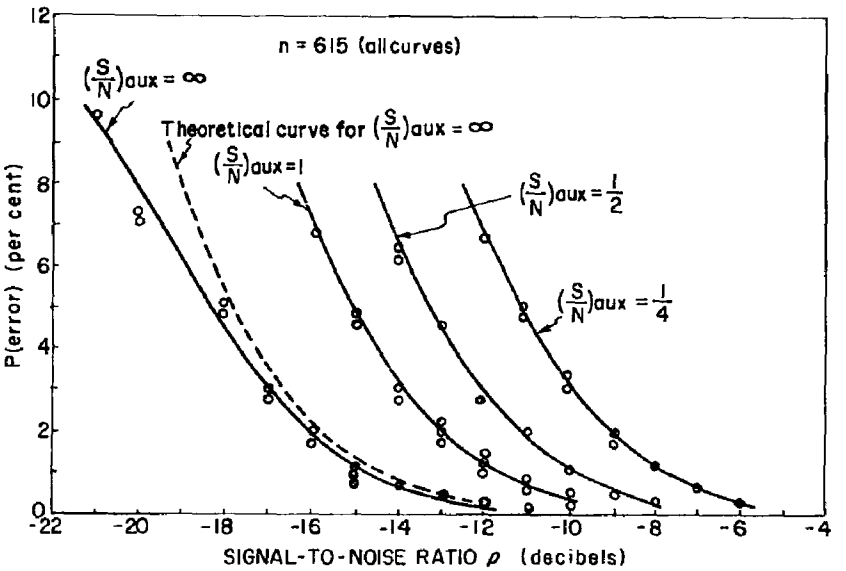

Fig. 7.6. Experimental probability of error vs. signal-to-noise ratio, wide-band integrating filter.

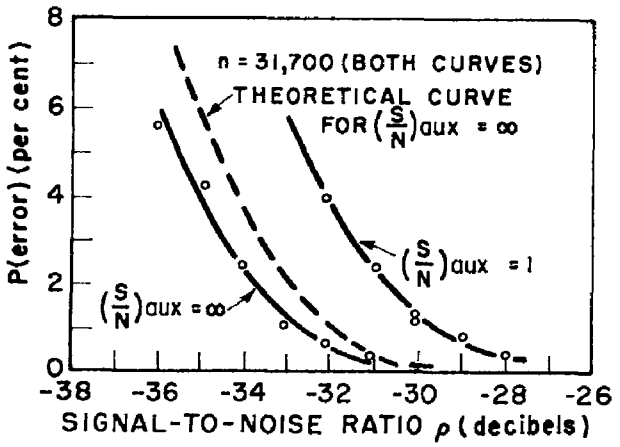

Fig. 7.7. Experimental probability of error vs. signal-to-noise ratio, narrow-band integrating filter.

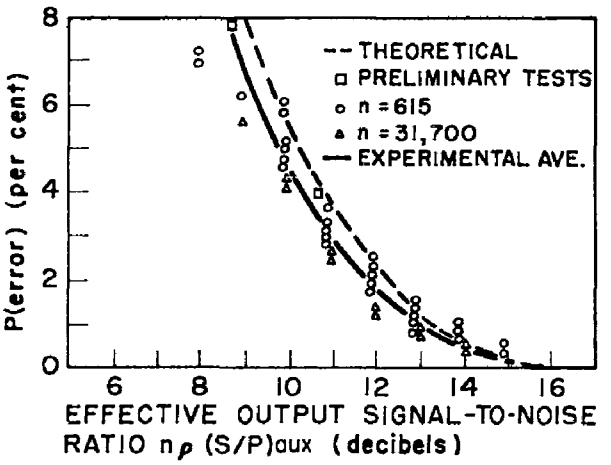

Fig. 7.8. Theoretical and experimental results of probability of error as a function of effective nρ.

A second source of error was that of reading the meters used in setting output levels. This process was made more difficult, since the output signals contained noise with significant low-frequency components.

The average of the observed results taken experimentally shows a shift to the left of approximately one-half a decibel. Perhaps the principal contribution to this shift (amounting to some 12 per cent) was the error in measuring the bandwidth of the input signals at the correlation receiver. In order to avoid such possible sources of error as saturation in amplifiers and Miller effect, the measurements were made at low level from the frequency-response pattern appearing on an oscilloscope. Under this condition, the noise output of the amplifier was comparable to the amplitude of the response pattern. Consequently, the amount of error apparently present does not appear to be inconsistent.

Still another source of error undoubtedly contributed to the spread of experimental results and to the relative increase of the probability of error for low values of probability. The presence of high-intensity radiation from various sources (e.g., radar) in and around the laboratory was observed to cause distinct errors in the counts made by the probability analyzer. An attempt to minimize these errors was made by taking counts at times when most other laboratory functions were closed down. The errors due to outside sources were kept small enough to be consistent with other experimental errors in the range over which measurements were taken, but they were increasingly significant at very low probability of error, and, in fact, set a practical lower limit to the range of observation.

It should be mentioned that each point plotted represents an average of several counts, and, where two or more points are plotted in Figs. 7.6 and 7.7 for a given condition, they represent readings taken on different days and under possibly different conditions of laboratory temperature, noise level, and equipment adjustment. They are plotted separately to indicate the reproducibility of the results.

CHAPTER VIII

RELATED TOPICS TO NOMAC SYSTEM DESIGN

A. The Effect of Nonideal Integration

The validity of the theoretical results thus far obtained is limited by the accuracy of the initial assumptions, as, for example, true Gaussian amplitude distributions, ideal short-time integration, distortionless multipliers, and rectangular pass bands. However, the effects of certain of these factor's being less perfect than assumed can be ascertained, and to some extent quantitative corrections can be made. Since experimenters have found that most natural noise sources (thermal agitation, shot noise, etc.) do have a Gaussian amplitude distribution at least for amplitudes less than about six times the standard deviation,[14] it seems evident that, if physical systems are operated in such a manner that saturation and/or cut-off effects may be neglected, the assumption of Gaussian noise seems well founded. Considerably larger differences may occur, however, if the integration method (filtering) is nonideal or if the multiplier used in the correlation process introduces appreciable distortion.

To investigate the effect of nonideal integration, it will be assumed that a correlation output W results from the sum of weighted components w_i. Here $w_i = x_i z_i$ is a product component of the type obtained when the functions X(t) and Z(t) are represented in the form indicated by Eq. (1-6). The results are limited to low-pass functions as derived, although the final result agrees with results obtained for band-pass functions when viewed from a different approach.[4]

The weight associated with each w_i is h_i, and the h's are constants and independent of the w's. Thus

$$W = \sum_{i=1}^{n} h_i w_i \quad . \tag{8-1}$$

Obviously, if all h_i equal unity, the W resulting is that discussed earlier under the assumption of ideal integration.

Now,

$$\overline{W} = E\left\{\sum_i h_i w_i\right\} = \sum_i E\left\{h_i w_i\right\} \quad , \tag{8-2}$$

$$= \overline{W} \sum_i h_i \quad . \tag{8-2a}$$

Similarly,

$$\overline{W^2} = E\left\{\sum_i h_i w_i \sum_j h_j w_j\right\} \quad , \tag{8-3}$$

$$= \sum_i \sum_j h_i h_j E\left\{w_i w_j\right\} \quad . \tag{8-3a}$$

But w_i is independent of w_j for $i \neq j$, so that one obtains

$$\overline{W^2} = \sum_i \sum_j h_i h_j \left\{\overline{w_i^2}\, \delta_j^i + (1 - \delta_j^i)\, \overline{w}_i^2\right\} \quad , \tag{8-4}$$

where δ_i^j is the Kronecker delta.

From the knowledge above of $\overline{W}$ and $\overline{W^2}$, the variance σ_W^2 may be obtained. Thus

$$\sigma_W^2 = \overline{W^2} - \overline{W}^2 = \sum_i \sum_j h_i h_j \left\{ \overline{w}_i^2 \delta_i^j + (1 - \delta_i^j)\, \overline{w}_i^2 \right\} - \overline{w_i}^2 \sum_i \sum_j h_i h_j \quad , \tag{8-5}$$

$$= [\overline{w}_i^2 - \overline{w}_i^2] \sum_i h_i^2 = \sigma_w^2 \Sigma h_i^2 \quad . \tag{8-6}$$

Again, Eq. (8-6) is seen to reduce to the case presented earlier when all h_i equal unity, i.e., ideal integration.

Since the errors in the systems considered here are caused by fluctuations about the average value of the output, the ratio of importance is

$$\frac{\sigma_W}{\overline{W}} = \frac{\sigma_w}{\overline{w}} \frac{\sqrt{\Sigma h_i^2}}{\Sigma h_i} = \frac{\sigma_w}{\sqrt{n}\,\overline{w}} \frac{\sqrt{\frac{1}{n}\Sigma h_i^2}}{\frac{1}{n}\Sigma h_i} \geq \frac{1}{\sqrt{n}} \frac{\sigma_w}{\overline{w}} \quad , \tag{8-7}$$

(because an rms value is never less than the average value). But the term on the right is the result obtained with ideal integration. Thus, there will be more frequent errors if nonideal integration is used.

To apply this result, the probability of error expression Eq. (4-30) is examined. Here

$$P\,(\text{error}) \sim \frac{K}{\sqrt{2\pi}} \sqrt{\frac{P}{nS}} \exp\left[-\frac{nS}{2P}\right] \quad .$$

In this expression S is σ_W^2 and the average $\overline{W}$ is given by $S\sqrt{n/P}$. If, instead, the values for σ_W^2 and $\overline{W}$ that would result from nonideal integration are substituted, the expression becomes

$$P(\text{error}) \sim \frac{K}{\sqrt{2\pi}} \sqrt{\frac{S}{P}} \frac{\sqrt{\Sigma h_i^2}}{\Sigma h_i} \exp\left[-\frac{S}{2P} \frac{(\Sigma h_i)^2}{\Sigma h_i^2}\right] \quad . \tag{8-8}$$

The ratio $(\Sigma h_i)^2/\Sigma h_i^2$ appears twice and by definition will be the effective n. Since n is twice the time-bandwidth product, the effective n is thus defined in terms of $2\tau W$ (this W is bandwidth), where τ is the effective integration time. It follows by definition that

$$\tau = \frac{1}{2W} \frac{\left(\sum_i h_i\right)^2}{\sum_i h_i^2} \quad . \tag{8-9}$$

If the approximations $nS/2N = C'T$ and $\ln K = H'T$, valid for $K \gg 1$, $n \gg 1$, and $S \ll N$ are used here, as in Chapter IV, the expression for probability of error becomes

$$P(\text{error}) \sim \frac{1}{2\sqrt{\pi\, C'\tau}} \exp\left[-C'\tau + H'T\right] \quad , \tag{8-10}$$

or

$$P(\text{error}) \sim \frac{1}{2\sqrt{\pi\,(\frac{\tau}{T})\, C'T}} \exp\left[-C'(\tfrac{\tau}{T})\, T \left(1 - \frac{H}{(\frac{\tau}{T})\, C'}\right)\right] \quad . \tag{8-10a}$$

Here $(\tau/T)C'$ appears in each place where C' appeared in Eq. (4-30). It is interpreted as a decrease of effective system capacity corresponding to the factor τ/T which will never exceed unity, since

$$\frac{\left(\sum_{i=1}^{n} h_i\right)^2}{\sum_{i=1}^{n} h_i^2}$$

is less than n. This in turn is true since the time devoted to signaling must be at least as long as the time devoted to averaging the correlated signal.

The integration time can be obtained from the impulse response function of the integration filter. One notes that the weights h_i themselves are values of the impulse response function at time intervals $\Delta t = 1/2w$. Therefore

$$\sum_i h_i = \sum_i h\left(\frac{i}{2w}\right) \tag{8-11}$$

where $h(t)$ is the filter impulse response. Furthermore, $\Sigma h\,(1/2w)$ is given approximately by

$$2W\int_o^T h(t)\,dt \quad .$$

Therefore, the effective integration time is given by

$$\tau \doteq \frac{\left[2W\int_o^T h(t)\,dt\right]^2}{(2W)^2\int_o^T h(t)^2\,dt} = \frac{\left[\int_o^T h(t)\,dt\right]^2}{\int_o^T h(t)^2\,dt} \quad . \tag{8-12}$$

Equation (8-12) is based on the assumption that $h(t)$ is zero for all t greater than T. This implies that $\tau < T$. Again, it is noted that, for an ideal integration function $h(t) = 1$ for $0 < t < T$, and zero elsewhere, $\tau = T$, and the result is the same as that obtained in Chapter IV.

Equation (8-12) may be manipulated in the following way: the difference between the integral with upper limit T and upper limit ∞ should be zero (or at least negligible). Then

$$\tau \doteq \frac{\left[\int_o^\infty h(t)\,dt\right]^2}{\int_o^\infty h(t)^2\,dt} = \frac{H(o)^2}{\frac{1}{2\pi}\int_{-\infty}^{\infty} |H(\omega)|^2\,d\omega} \quad , \tag{8-13}$$

or

$$\doteq \frac{H(o)^2}{\int_{-\infty}^{\infty} |H(f)|^2\,df} \quad . \tag{8-13a}$$

This is obviously the reciprocal of what is defined as the effective or noise bandwidth. Thus

$$\tau \doteq \frac{1}{\Delta f} \quad . \tag{8-14}$$

The noise bandwidth Δf is defined for double-sided spectra (defined for $-\infty < f < \infty$). Thus, the input noise bandwidth, since rectangular channel pass bands are assumed, is $2W$. This means

$$n = \frac{2W}{\Delta f} \quad , \tag{8-14a}$$

the ratio of channel noise bandwidth to integrating filter noise bandwidth. This result is obtained from a different approach in Ref. 4.

B. The Effect of Distortion in Multipliers

The correlation process inherently implies multiplication, yet electronic multiplication of two time-varying parameters is not easily accomplished. Probably the simplest type of electronic multiplication is that which takes place in the conventional converter stage of a superheterodyne receiver, and is the type used in the correlation converter of the experimental system.

The first-order analysis of tubes of the type used as multipliers is that the output current is the product of two quantities, $a + f_1(t)$ and $b + f_2(t)$. The result is, of course, $ab + bf_1(t) + af_2(t) + f_1(t)f_2(t)$. The constants a and b as well as the components $bf_1(t)$ and $af_2(t)$ may be separated from the product $f_1(t)f_2(t)$ by band-pass filtering, if $f_1(t)$ and $f_2(t)$ are functions occupying distinct frequency bands. Thus, in view of the elementary analysis, the converter-tube multipliers are as good as the filtering used to separate the desired component from the undesired ones.

However, a more general approach considers the effects of distortion in the multiplying process. In general, the time function of the output may be expressed by the summation

$$\psi(t) = \sum_{i=o}^{\infty} \sum_{j=o}^{\infty} a_{ij} f_1{}^{i}(t) f_2{}^{j}(t) \quad . \tag{8-15}$$

The $\psi(t)$ thus obtained is, for example, the plate current of the converter tube to which $f_1(t)$ is applied at one grid and $f_2(t)$ is applied at another. It is further assumed that the spectra of $f_1(t)$ and $f_2(t)$ do not overlap. The term yielding the desired output is that associated with a_{11}, and in particular, one of the two side bands – say, that at the difference frequency – is isolated by filtering. Thus, the desired output is

$$\psi_d(t) = \frac{a_{11}}{2} f_1(t)\, f_2(t) \quad . \tag{8-16}$$

The products of distortion that lie in the product-frequency band are the only ones that will lead to errors in the results. How these occur and their relative significance can be obtained by considering as typical functions those defined by

$$f_1(t) = x_1 \cos \omega_1 t; \; f_2(t) = x_2 \cos \omega_2 t \quad . \tag{8-17}$$

Then

$$\left[\cos \omega_1 t\right]^r = \frac{1}{2^r} \exp\left[-jr\omega_1 t\right] \sum_{k=o}^{r} {}_rC_k \exp\left[2jk\omega_1 t\right] \quad , \tag{8-18}$$

and

$$\left[\cos \omega_2 t\right]^s = \frac{1}{2^s} \exp.\left[-js\omega_2 t\right] \sum_{h=o}^{s} {}_sC_h \exp\left[2jh\omega_2 t\right] \quad . \tag{8-18a}$$

From these,

$$\left[\cos\omega_1 t\right]^r \left[\cos\omega_2 t\right]^s = \frac{1}{2^{r+s}} \sum_{k=o}^{r} \sum_{h=o}^{s} {}_rC_k \, {}_sC_h \exp\left[jt\,(2k-r)\,\omega_1 + (2k-s)\,\omega_2\right] \quad . \tag{8-18b}$$

Now, if the filtering is such that only the component at frequency $\omega_1 - \omega_2$ lies within the pass band, only those terms for which $2k - r = 1$ and $2h - s = -1$ will be passed by the filter. For these conditions,

$$k = \frac{r+1}{2} \quad ; \qquad h = \frac{s-1}{2} \quad . \tag{8-19}$$

Therefore, the component at the difference frequency that will appear at the output of the filter will be

$$\frac{{}_rC_{(r+1)/2} \; {}_sC_{(s-1)/2}}{2^{r+s-1}} \cos\,(\omega_1 - \omega_2)\,t \quad . \tag{8-20}$$

It is noted that $(r + 1)/2$ and $(s - 1)/2$ must be integers, and that consequently r and s must be odd integers. Furthermore, because of the symmetry of the binomial expansion, ${}_sC_{(s-1)/2} = {}_sC_{(s+1)/2}$ and the latter may be substituted in the given expression for greater symmetry. Thus, where the functions $\cos \omega_1 t$ and $\cos \omega_2 t$ differ by a third frequency not harmonically related to either of the other two, the expression for the portion of the product function appearing at the output of the filter is

$$\psi_o(t) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} \frac{{}_iC_{(i+1)/2} \; {}_jC_{(j+1)/2}}{2^{i+j-1}} x_1^i x_2^j \cos\,(\omega_1 - \omega_2)\,t \quad . \tag{8-21}$$

Here x_1 and x_2 are the constants by which sinusoidal terms $\cos \omega_1 t$ and $\cos \omega_2 t$ are multiplied, and the expression (8-21) shows the contribution of distortion terms in a converter-multiplier using filtering in the output.

As thus derived, x_1 and x_2 may be thought of as quasi-stationary coefficients – that is, x_1 and x_2 may vary with time if those variations are slow compared to the frequency representing the filter bandwidth. Obviously, this leads to contradictions if the signals being analyzed are limited in their fluctuations by the response of the filter. However, the exact analysis is sufficiently more difficult to justify using expressions such as (8-21) as a guide to the effects of multiplier distortion.

It is noted in passing that the coefficients a_{ij} are related to the coefficients of a Taylor expansion by the expression

$$a_{ij} = \frac{1}{i!j!} \left[\frac{\partial^{i+j} \psi(t)}{\partial f_1(t)^i \, \partial f_2(t)^j} \right]_{t=t_o} \quad , \tag{8-22}$$

from which one obtains

$$\psi\left[f_1(t_o + t), \, f_2(t_o + t)\right] = a_{oo} + a_{o1} f_2(t) + a_{1o} f_1(t) + \ldots \quad . \tag{8-23}$$

Experiments conducted by R. S. Berg and B. Eisenstadt (as yet unpublished) have

shown that, for a reasonable choice of tube operating point, most converter tubes tested had contributions due to a_{31} and a_{13} terms 40db or more below the desired a_{11} terms. Higher-order terms were found to be appreciably smaller.

Therefore, it is found that $(3/8)a_{31} < .01\,(1/2\,a_{11})$, where the coefficients 3/8 and 1/2 are the values obtained from Eq. (8-20). The numbers a_{ij} are, of course, those that might be obtained from an exhaustive analysis of the static characteristics of the converter tube. It is further noted that the effective values of the terms of 6th order are $(5/32)a_{51}$, $(5/32)a_{15}$ and $(9/32)a_{33}$. Since the a's themselves represent 6th order derivatives of rather smooth tube characteristics, they are quite small, and will be neglected along with all higher-order terms in the discussion that follows.

If it is assumed that the a_{31} and a_{13} terms are of about the same magnitude, it follows that the average output W is given by

$$\overline{W} \doteq E\left\{\sum_{i=1}^{n}\left[X_iY_i + \frac{3}{4}a_{13}X_iY_i^3 + \frac{3}{4}a_{31}X_i^3Y_i\right]\right\} \quad . \tag{8-24}$$

From this, it is seen that, for x and y independent, $\overline{W}$ is zero, but for x and y not independent, the average output of the multiplier is increased by the distortion terms and, in particular, if $x = y$, the increase is about $d \cdot x^4$. The d used here is the measured distortion coefficient $(a_{31} + a_{13})$.

A case of greater interest, however, is that arising when y is the sum of an independent noise and a part of x. Thus, y becomes $y + ax$. Here the ratio of $\overline{(ax)^2}$ to $\overline{Y^2}$ is the familiar S/N or ρ.

If the magnitudes of x and y are about the same, as is desired in practical cases, the substitution of these values into Eq. (8-24) (for $\rho \ll 1$) will yield

$$\overline{W} = n\sqrt{S\,\overline{e_g^2}}\,\left[1 + 6d\,\overline{e_g^2}\,\left(1 + \frac{1}{2}\rho\right)\right] \quad . \tag{8-25}$$

where $\overline{e_g^2}$ is the mean-square grid signal voltage applied to the tube.

In a similar manner, involving a good deal of algebra that is omitted here, the variance of W is found to be

$$\sigma_W^2 \doteq n\,\overline{e_g^2}\,N\left[1 + 12d\,\overline{e_g^2}\,\left(1 + \frac{3}{2}\rho\right)\right] \quad . \tag{8-26}$$

The ratio of the standard deviation to average output in the absence of multiplier distortion is found to be $\sqrt{N/nS}$ or $(E_s/N_o)^{-\frac{1}{2}}$. Here, it is seen that the ratio becomes

$$\left(\frac{\sigma_W}{\overline{W}}\right)_{\text{actual}} \doteq \left(\frac{\sigma_W}{\overline{W}}\right)_{\text{ideal}} \frac{\left[1 + 12d\,\overline{e_g^2}\,\left(1 + \frac{3}{2}\rho\right)\right]^{\frac{1}{2}}}{\left[1 + 6d\,\overline{e_g^2}\,\left(1 + \frac{1}{2}\rho\right)\right]} \quad , \tag{8-27}$$

in which $\rho \ll 1$ and $\overline{e_g^2}$ is less than unity. Under these conditions, the effect of multiplier distortion is kept quite small, and may be ignored in the final results, at least where the assumption of reasonable operating point and attendant value of d about 0.01 is valid.

CHAPTER IX

CONCLUSIONS

The results thus far given are summarized as follows. From the theoretical point of view, it has been shown that NOMAC systems are members of a class of optimum communication systems when the noise is considerably larger than the signal. NOMAC systems deliver to the user of the communicated information a choice made by the receiver from the set of possible transmitted symbols. The class is optimum in that the symbol thus chosen is, with probability approaching unity, the most probable (a posteriori) of the symbols that might have been transmitted.

The fundamental parameter limiting the performance of these systems is the ratio of signal energy to noise power per cycle. Any combination of bandwidth ratio (or n = 2TW) and signal-to-noise ratio leading to the same ratio Es/N_o is capable of substantially the same performance, within the limits of signal-to-noise ratio less than unity and bandwidth ratio much greater than unity. In particular, small signal-to-noise ratio may be offset with large bandwidth ratio, regardless of how small the former might become.

It is shown that, if the decision made by the receiver is based on maximum correlation, for information rates not exceeding the channel capacity, the per-unit equivocation may be reduced to any desired value by allowing sufficient delay (coding time).

On the other hand, if the detection criterion is that of exceeding a threshold, an equivalent loss of channel capacity is occasioned by an arbitrary choice of the threshold. Even when the optimum threshold is used, a delay an order of magnitude greater than that required when the criterion of detection is maximum correlation is necessary to obtain the same per-unit equivocation. The compensation in favor of threshold detection is the reduction of equipment complexity in the receiver performing this type of detection.

An important disadvantage is connected with the results just given. This disadvantage is the stringent requirements of synchronization and storage associated with ideal cross-correlation, in which process copies of each of the possible message waveforms are available for correlation at the receiver. If it is assumed that storage facilities do exist at the receiver capable of storing the possible signals, the ease of using random samples of noise to represent the transmitted symbols is lost, and the problem of time synchronization to enable correlation to be performed at the "$\tau = o$" point of the correlation curve is one of major proportions.

It is shown that, if transmission of the reference copies of the possible signals as well as the intelligence signal is resorted to, there is a loss of effective channel capacity (or effective ratio of Es/N_o) corresponding to the ratio of signal power to signal-plus-noise power in the auxiliary channel or channels. This use of auxiliary channels implies, in addition, an expenditure of at least twice the bandwidth used by stored signal systems. To its advantage, the so-called "two-signal system" is free of the synchronization problem, and makes possible the use of random samples of currently generated noise as symbols.

An additional apparent disadvantage is the "self-noise" that results from the short-time auto-correlation of noise signals. It is only an apparent fault, however, in that this

component of noise is (1) small compared to other noise for small values of channel signal-to-noise ratio, and (2) precisely specified by the knowledge of the waveform being correlated (which is assumed) and thus could (with sufficient equipment complexity) be eliminated.

One version of the stored signal type of NOMAC system has been suggested by Fano.[19] It is a variation of the matched-filter technique of North, Middleton and Van Vleck, and seems sufficiently promising to merit further investigation. This system apparently solves the synchronization and storage problems simultaneously at the expense of somewhat greater per-unit equivocation. A complete introduction to the method is found in the reference cited.[19]

From the practical point of view, it is demonstrated that laboratory models of NOMAC are easily realized in comparatively compact equipment. This is partly a result of the use of converter tubes as multipliers, which are shown to contribute insignificant error when properly used. Also, the use of simple filters as integrators is shown to be feasible. While the per-unit equivocation is increased by the reduction in effective integration time (delay) when filters are used to perform integration, it is felt that the disadvantage is compensated by simplicity of equipment.

Much remains to be determined about the properties of NOMAC and related communication systems. It is recalled that all the results reported here, theoretical and experimental (as far as practicable), are given for but one type of channel disturbance – that of additive white Gaussian noise. How these systems will behave in the presence of other noise effects, or when they are subject to multipath propagation, remains an important problem for further investigation.

Also, a detailed study of combinations of various modulation methods and random carrier signals is recommended. For example, a frequency-modulated random carrier* has been tried in the laboratory with considerable success, and obviously merits extensive investigation to determine its applicability to secure voice-communication links.

*Suggested by R. M. Fano, 11 June 1951.

REFERENCES

1. J. R. Carson, Notes of the Theory of Modulation, Proc. IRE, Feb. 1922.
2. G. A. Campbell and R. M. Foster, Fourier Integrals for Practical Applications, Bell Tel. Lab. Series, Van Nostrand, New York, 1948.
3. W. B. Davenport, Jr., A Study of Speech Probability Distributions, M.I.T. – R.L.E. Technical Report #148, Aug. 1950.
4. W. B. Davenport, Jr., Correlator Errors Due to Finite Observation Intervals, M.I.T. – R.L.E. Technical Report #191, Mar. 1951.
5. R. M. Fano, Transmission of Information I&II, M.I.T. – R.L.E. Technical Report #65 and #149, Mar. 1949 and Feb. 1950.
6. R. M. Fano, Short-Time Correlation Functions and Power Spectra, Jour. Acous. Soc. Amer., Sept. 1950.
7. R. M. Fano, On Signal-to-Noise Ratios in Correlation Detectors, M.I.T. – R.L.E. Technical Report #186, Feb. 1951.
8. P. Franklin, Methods of Advanced Calculus, McGraw Hill, New York, 1944.
9. T. C. Fry, Probability and its Engineering Uses, Van Nostrand, New York, 1946.
10. D. Gabor, Communication Theory and Physics, Phil Mag., Nov. 1950.
11. M. J. Golay, Note on the Theoretical Efficiency of Information Reception with PPM, Proc. IRE, Sept. 1949.
12. R. V. L. Hartley, Transmission of Information, B.S.T.J., July 1948.
13. F. B. Hildebrand, Advanced Calculus for Engineers, Prentiss-Hall, New York, 1949.
14. N. Knudsen, An Experimental Study of Random Noise, M.I.T. – R.L.E. Technical Report #115, July 1949.
15. Y. W. Lee, T. P. Cheatham, Jr., J. B. Wiesner, The Application of Correlation Functions in the Detection of Small Signals in Noise, M.I.T. – R.L.E. Technical Report #141, Oct. 1949.
16. Y. W. Lee, Application of Statistical Methods to Communication Problems, M.I.T. – R.L.E. Technical Report #181, Sept. 1950.
17. H. Nyquist, Certain Factors Affecting Telegraph Speed, B.S.T.J., April 1924.
18. E. Reich, On the Definition of Information, Jour. Math. & Phys. Oct. 1951.
19. Research Laboratory of Electronics – PROJECT LINCOLN, Quarterly Progress Report, (Classified) 30 Jan. 1952.
20. S. O. Rice, Communication in the Presence of Noise - Probability of Error for Two Encoding Schemes, B.S.T.J., Jan. 1950.
21. C. E. Shannon, Communication in the Presence of Noise, Proc. IRE, Jan. 1949.
22. C. E. Shannon and W. Weaver, The Mathematical Theory of Information, University of Illinois Press, 1949.
23. W. G. Tuller, Theoretical Limitations on the Rate of Transmission of Information, Proc. IRE, May 1949.
24. N. Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series, John Wiley, New York, 1949.
25. N. Wiener, Cybernetics, John Wiley, New York, 1948.
26. P. M. Woodward, and I. L. Davies, Information Theory and Inverse Probability in Telecommunications, Telecomm. Res. Est. Tech. Note #137, Sept. 1951, Proc. IRE, March 1952.

ACKNOWLEDGEMENTS

This thesis investigation was performed in the Research Laboratory of Electronics (and later in Division 3 of PROJECT LINCOLN). Throughout the work the author received the close cooperation of and helpful suggestions from the members of the Laboratory staff. For their time and assistance during the course of the investigation, the author would like to express his thanks, in particular to Professors R. M. Fano, W. B. Davenport, Jr., and J. B. Wiesner; for their suggestions and contributions which saved the author much effort, he would like to thank, among others, Messrs. P. Green, Jr., W. C. McLaughlin, R. Price, and Mr. T. Sarantos, the technician responsible for most of the experimental equipment.

APPENDIX I

PER-UNIT EQUIVOCATION AND PROBABILITY OF ERROR

If H_x represents the entropy of the information source and $H_{x/z}$ is the equivocation or remaining uncertainty inherent in the channel, the per-unit equivocation is defined as

$$\text{P.U.E.} = \frac{H_{x/z}}{H_x} \quad . \tag{A1-1}$$

For systems of the type considered here, $H_x = \log (K + 1)$. The equivocation is given (see Ref. 5) by

$$H_{x/z} = -\sum_z P(z) \sum_z P(x/z) \log P(x/z) \quad . \tag{A1-2}$$

Expression (A1-2) is interpreted to mean the average over-all received signals of the uncertainty represented by the a posteriori probability distribution following reception of a signal. For the purposes of calculation, the a priori probability of any one of the possible symbols is taken as $1/K + 1$. The a posteriori probability of the indicated symbol is $(1 - p)$, where p is the probability of error, and the a posteriori probability for each of the remaining K symbols is taken as p/K. The equivocation then becomes

$$H_{x/z} = -\left[(1 - p) \log (1 - p) + p \log \frac{p}{K}\right] \quad . \tag{A1-3}$$

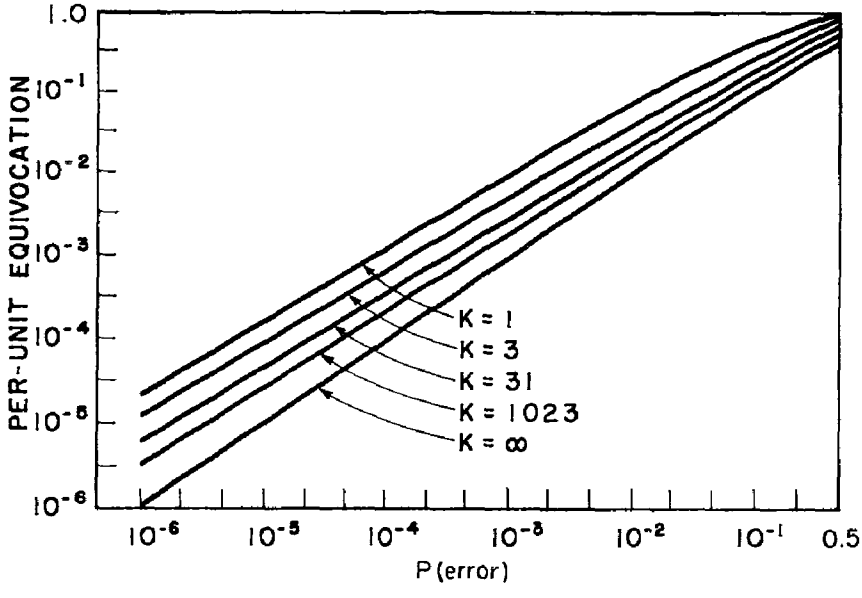

Fig. A1.1. Per-unit equivocation vs. probability of error.

The resulting expression for the per-unit equivocation, which is the upper limit that can be associated with any given probability of error, is then

$$\text{P.U.E.} = \frac{(p-1)\log(1-p) - p\log\frac{p}{K}}{\log K + 1} \quad . \tag{A1-4}$$

An expression for P.U.E. when p becomes very small (the practical case) is as follows.

$$\text{P.U.E.} \doteq p\left\{\frac{\log e + \log\frac{K}{p}}{\log K + 1}\right\} \quad . \tag{A1-5a}$$

Plotted in Fig. A1.1 and tabulated below are typical relations between P.U.E., p, and K.

TABLE AI-1

PER-UNIT EQUIVOCATION
Corresponding To Probability Of Error When Receiver Chooses
Most Probable Message As Message Transmitted

K + 1	2	4	32	128	1024	1048576	
log K + 1	1	2	5	7	10	20	
p							
0.5	1.0	0.877	0.695	0.643	0.600	0.555	0.500
10^{-1}	4.7p	3.14p	1.93p	1.67p	1.47p	1.235p	p
10^{-2}	8.1p	4.84p	2.61p	2.15p	1.81p	1.41p	p
10^{-3}	11.04p	6.31p	3.20p	2.57p	2.10p	1.55p	p
10^{-4}	14.72p	8.15p	3.93p	3.10p	2.47p	1.74p	p
10^{-5}	18.00p	9.81p	4.60p	3.58p	2.80p	1.90p	p
10^{-6}	21.30p	11.45p	5.26p	4.05p	3.14p	2.07p	p

APPENDIX II

THE DISTRIBUTION OF VECTOR MAGNITUDES

Given: the probability density distribution for x, namely,

$$p(x) = \frac{1}{\sqrt{2\pi S}} \exp\left[-\frac{x^2}{2S}\right] \quad . \tag{A2-1}$$

Now $|X|^2 = \sum_i x_i^2$, so that

$$p(x_i^2) = \frac{\exp\left[-\frac{x_i^2}{2S}\right]}{2\sqrt{x_i^2}\sqrt{2\pi S}} \quad , \tag{A2-2}$$

and, where $\phi(t)$ is the characteristic function related to the distribution,

$$\phi'(t) = \int_{-\infty}^{\infty} \exp[itx_i^2]\, p(x_i^2)\, dx_i^2 \quad , \tag{A2-3}$$

$$= [1 - 2iSt]^{-\frac{1}{2}} \quad , \tag{A2-4}$$

then

$$\phi_o'(t) = [\phi'(t)]^n = [1 - 2iSt]^{-\frac{n}{2}} \quad , \tag{A2-5}$$

and

$$p_o'(|X|^2) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp[-it|X|^2]\, \phi_o'(t)\, dt \quad . \tag{A2-6}$$

$$p_o(|X|) = \frac{2|X|^{n-1} \exp\left[-\frac{|X|^2}{2S}\right]}{(2S)^{\frac{n}{2}}\,\Gamma(\frac{n}{2})} \quad . \tag{A2-7}$$

The distribution (A2-7) may be derived in another way which reveals significantly what is involved in $p_o(|X|)$.

First, the probability density distribution for a given point in n-space is given by the product of the probability density distribution for each of the Cartesian coordinates (the coordinates are independent). Thus

$$p''(X) = p(X_1)\, p(X_2) \ldots\ldots p(X_n) \quad , \tag{A2-8}$$

$$= \left(\frac{1}{2\pi S}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2S}(X_1^2 + X_2^2 + \ldots . X_n^2)\right] \quad , \tag{A2-9}$$

$$= \left(\frac{1}{2\pi S}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2S}(|X|^2)\right] \quad . \tag{A2-10}$$

In (A2-10), $|X|^2$ is the squared magnitude of the distance of a point X from the origin, and the probability density of the point is seen to depend only on this distance.

The probability density distribution for the distance $|X|$ is the integral of $p''(X)$ taken over all points that lie distance $|X|$ from the origin. Since $|X|$ is constant in the integration, $p_o(|X|)$ is the surface area ($n-1$ space) of the hypersphere of radius $|X|$ times $p''(X)$. The surface area is in turn the rate of change of the volume of the hypersphere with respect to the radius at that value of radius.

The volume of a hypersphere in n-space can be obtained as follows.

$$V_n(|X|) = \int_V dV = 2^n \int_o^{|X|} d\zeta_1' \int_o^{\sqrt{|X|^2 - \zeta_1'^2}} d\zeta_2' \cdots \int_o^{\sqrt{|X|^2 - \sum_{i=1}^{n-1} \zeta_i'^2}} d\zeta_n' \quad . \tag{A2-11}$$

Let

$$\sigma_j^2 = |X|^2 - \sum_{i=1}^{n-j} \zeta_i'^2 \quad ,$$

then

$$\sigma_n^2 = |X|^2 \quad ,$$

and

$$V_n(|X|) = 2^n \int_o^{\sigma_n} d\zeta_1' \int_o^{\sigma_{n-1}} d\zeta_2' \cdots \int_o^{\sigma_2} d\zeta_{n-1}' \int_o^{\sigma_1} d\zeta_n' \quad , \tag{A2-12}$$

$$V_n(|X|) = 2^n \int_o^{\sigma_n} d\zeta_1' \int_o^{\sigma_{n-1}} d\zeta_2' \cdots \int_o^{\sigma_2} \sqrt{\sigma_2^2 - \zeta_{n-1}'^2}\, d\zeta_{n-1}' \quad . \tag{A2-13}$$

But

$$\int_o^{\sigma_2} \left(\sigma_2^2 - \zeta_{n-1}'^2\right)^{\frac{1}{2}} d\zeta_{n-1}'$$

is recognized as the Beta Function $(1/2\,\sigma_2^2)\,\beta(1/2, 3/2)$. Generally

$$\int_o^a (a^2 - \rho^2)^{q/2}\, d\rho = \tfrac{1}{2} a^{\frac{q+3}{2}} \beta\left(\tfrac{1}{2}, \tfrac{q+2}{2}\right) \quad .$$

Therefore,

$$V_n(|X|) = 2^{n-r} \frac{\Gamma(\frac{1}{2})^r \Gamma(\frac{3}{2})}{\Gamma(\frac{r+3}{2})} \int_o^{\sigma_n} d\zeta_1' \int_o^{\sigma_{n-1}} d\zeta_2' \cdots \int_o^{\sigma_{r+2}} \left(\sigma_{r+1}^2\right)^{r+1/2} d\zeta_{n-r-1}' \quad , \tag{A2-14}$$

so that when carried out to $r = n-1$,

$$V_n(|X|) = 2 \frac{\Gamma(\frac{1}{2})^{n-1} \Gamma(\frac{3}{2})}{\Gamma(\frac{n+2}{2})} \sigma_n^n \quad , \tag{A2-15}$$

$$= \frac{\pi^{n/2} |X|^n}{\frac{n}{2}\Gamma(\frac{n}{2})} \quad . \tag{A2-16}$$

The surface $S_n(|X|)$ of the hypersphere is thus readily obtained as

$$S_n(|X|) = \frac{2\pi^{n/2} |X|^{n-1}}{\Gamma(\frac{n}{2})} \quad , \tag{A2-17}$$

and the product of (A2-10) and (A2-17) is

$$p_o(|X|) = \frac{2\pi^{\frac{n}{2}} |X|^{n-1}}{\Gamma(\frac{n}{2})} \cdot \frac{\exp\left[-\frac{|X|^2}{2S}\right]}{(2\pi S)^{\frac{n}{2}}} \quad , \tag{A2-18}$$

$$= \frac{2|X|^{n-1} \exp\left[-\frac{|X|^2}{2S}\right]}{(2S)^{n/2}\Gamma(\frac{n}{2})} \quad . \tag{A2-7}$$

Some of the properties of this distribution are as follows.

$$E\left[|X|\right] = \int_o^\infty \frac{2|X|^n \exp\left[-\frac{|X|^2}{2S}\right]}{(2S)^{n/2}\Gamma(\frac{n}{2})} d|X| = \frac{1}{(2S)^{\frac{n}{2}}\Gamma(\frac{n}{2})}\left[(2S)^{\frac{n+1}{2}}\Gamma(\frac{n+1}{2})\right] \quad , \tag{A2-19}$$

$$= \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\sqrt{2S} \quad . \tag{A2-20}$$

$$\lim_{n\to\infty} E\left[|X|\right] = \sqrt{nS} \quad . \tag{A2-21}$$

$$E\left[|X|^2\right] = \int_o^\infty \frac{2|X|^{n+1} \exp\left[-\frac{|X|^2}{2S}\right]}{(2S)^{n/2}\Gamma(\frac{n}{2})} d|X| = \frac{1}{(2S)^{n/2}\Gamma(\frac{n}{2})}\left[(2S)^{\frac{n}{2}+1}\Gamma(\frac{n}{2}+1)\right] \quad , \tag{A2-22}$$

$$= nS \quad . \tag{A2-22a}$$

$$\sigma^2_{|X|} = nS - \left[\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}\sqrt{2S}\right]^2 \lesssim \frac{S}{2} \quad . \tag{A2-23}$$

If Stirling's approximation is used for $\Gamma(n)$ and one examines $p_o(|X|/\sqrt{nS})$, there results the expression

$$p_o\left(\frac{|X|}{\sqrt{nS}}\right) = \sqrt{\frac{n}{\pi}}\left(\frac{|X|}{\sqrt{nS}}\right)^{n-1} \exp\left\{-\frac{n}{2}\left[\left(\frac{|X|}{\sqrt{nS}}\right)^2 - 1\right]\right\} \quad , \tag{A2-24}$$

which is valid for n of the order of ten or higher.

One notes that, for $|X| = \sqrt{nS}$, the value of the density distribution function is $\sqrt{n/\pi}$ and that the variance of the distribution behaves as $1/n$. Therefore, the distribution function behaves somewhat like a unit impulse function at $|X|/\sqrt{nS} = 1$. The distribution is given for a few values of n, tabulated in Table AII-1, and plotted in Fig. A2.1.

TABLE AII-1

NORMALIZED PROBABILITY DENSITY OF VECTOR MESSAGE

n = 3		n = 10		n = 100	
$\left(\frac{\lvert X\rvert}{\sqrt{nS}}\right)^2$	$p_o\left(\frac{\lvert X\rvert}{\sqrt{nS}}\right)$	$\left(\frac{\lvert X\rvert}{\sqrt{nS}}\right)^2$	$p_o\left(\frac{\lvert X\rvert}{\sqrt{nS}}\right)$	$\left(\frac{\lvert X\rvert}{\sqrt{nS}}\right)^2$	$p_o\left(\frac{\lvert X\rvert}{\sqrt{nS}}\right)$
0.01	0.043	0.50	0.965	0.81	2.26
0.25	0.805	0.90	1.830	0.90	4.52
0.667	1.070	1.00	1.783	0.96	5.52
1.000	0.977	1.20	1.490	0.99	5.66
2.000	0.438	2.00	0.271	1.04	5.35
4.000	0.043			1.10	4.29
				1.21	1.96

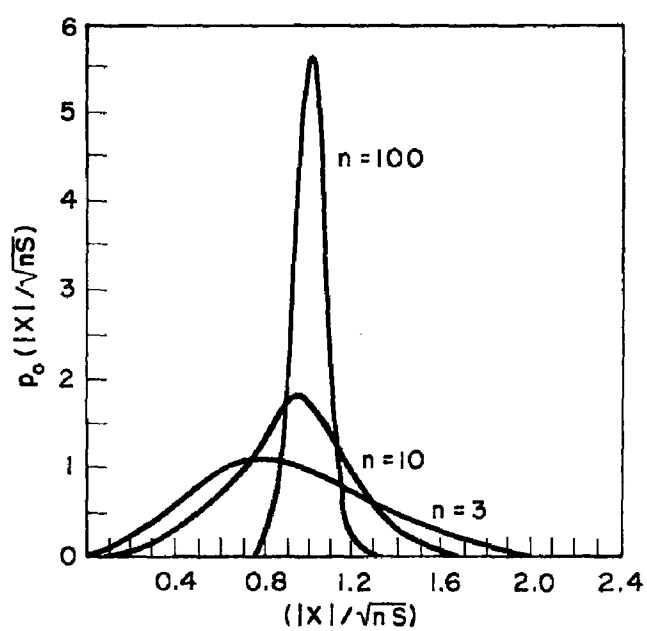

Fig. A2.1. Probability distribution of relative amplitudes of vectors.

APPENDIX III

PROBABILITY OF ERROR: ADAPTATION FROM S. O. RICE

Equation (4-12) of S. O. Rice's article "Communication in the Presence of Noise – Probability of Error for Two Encoding Schemes" is given (in the nomenclature of this paper) as

$$P(\text{no error}) = P(D_1, D_2, D_3, \cdots D_k > D_o) \quad ,$$

$$= \int_o^\infty d|Z| \int_o^\infty d|Y|\ p(|Z|, |Y|)\ \left[1 - P(|Y|, |Z|)\right]^K \quad . \tag{A3-1}$$

In his Appendix I, Rice gives as an approximation for $P(|Y|, |Z|)$ the expression

$$P(|Y|, |Z|) \sim \frac{1}{\sqrt{2\pi}} \left(\frac{\sigma_D}{\mu}\right) \exp\left[-\frac{\mu^2}{2\sigma_D^2}\right] \quad , \tag{A3-2}$$

where $\sigma_D^2 = E\left[D - \vec{D}\right]^2$ and $\mu = \overline{D} - |Y|$. This expresses the probability that any one distance ("D" above) from the point Z to the point X_k be less than the magnitude of the noise vector $\vec{Y}$, under conditions of fixed $|Z|$ and fixed $|Y|$. The result is then averaged over all magnitudes of $\vec{Z}$ and $\vec{Y}$.

An approximation for the probability of error results by taking $1 - P(D_1, D_2, D_3, \cdots D_k > D_o)$ which for small P(error) is given approximately by

$$P(\text{error}) \doteq K \int_o^\infty d|Z| \int_o^\infty d|Y|\ p(|Z|, |Y|)\ P(|Y|, |Z|) \quad . \tag{A3-3}$$

As n increases, $p(|Z|, |Y|)$ approaches an impulse with all significant contribution to the integral in the vicinity of $|Z| = nP$ and $|Y| = nN$. Evaluating the integral at these values of $|Z|$ and $|Y|$ yields a rough approximation for the probability of error, namely,

$$P(\text{error}) \sim \frac{K}{\sqrt{2\pi}} \left(\frac{\overline{\sigma_D}}{\overline{\mu}}\right) \exp\left[-\frac{\overline{\mu}^2}{2\overline{\sigma_D}^2}\right] \tag{A3-4}$$

The variance σ_D^2 defined earlier as the mean square difference of the distances D_k and their mean is given by Rice in terms of $|Z|$, but when evaluated at the average $|Z|$ equals $2nS(S + 2P)$. Rice gives μ in terms of $|Y|$. Again, if evaluated at the average $|Y|$, the result is $\overline{\mu} = \overline{D} - \overline{|Y|} = n(P + S - N) = 2nS$. For these values, the approximation reads

$$P(\text{error}) \sim \frac{K}{\sqrt{2\pi}} \left(\sqrt{\frac{S + 2P}{2nS}}\right) \exp\left[-\frac{nS}{S + 2P}\right] \quad . \tag{A3-5}$$

Where $S \ll N$, this may be written even more approximately as

$$P(\text{error}) \sim \frac{K}{\sqrt{2\pi}} \sqrt{\frac{N}{nS}} \exp\left[-\frac{nS}{2N}\right] . \tag{A3-6}$$

This expression can now be written in terms of the approximations $nS/2N \doteq C'T$ for $S \ll N$ and $K \doteq \exp[H'T]$ for $K \gg 1$. The result is then

$$P(\text{error}) \sim \frac{1}{2\sqrt{\pi C'T}} \exp\left[-C'T\left(1 - \frac{H}{C}\right)\right] , \qquad \text{(A3-7)}$$

and is reasonably valid for very large n, large K, and large noise-to-signal ratios. The first two conditions are the same as those introduced by Rice for the validity of his probability-of-no-error expression. The third is added to show the similarity of his results to those of Chapter IV.

APPENDIX IV

THE ARBITRARY ORIENTATION OF THE VECTOR FIELD

Let the set of possible message vectors be set up in a matrix $\left[S_{k+1,n}\right]$. The subscripts indicate rows and columns in that order, where the vectors are arrayed by rows, their coordinate components by columns.

$$\left[S_{k+1,n}\right] = \begin{bmatrix} X_{01} & X_{02} & \cdots & X_{0n} \\ X_{11} & X_{12} & \cdots & X_{1n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ X_{k1} & X_{k2} & \cdots & X_{kn} \end{bmatrix} \quad . \tag{A4-1}$$

It is assumed that this matrix is known to both the transmitter and receiver, and the transmission of a message is represented mathematically as the selection of one of the K + 1 rows (arbitrarily, the $\vec{X}_0$ vector), and adding another row matrix $\underline{Y}$ to this row and delivering the result to the receiver. The resulting row matrix is $\underline{X_0} + \underline{Y} = \underline{X_0 + Y}$. The transpose of this matrix is designated by $Z\rceil$, and at the receiver the operation of $\left[S_{k+1,n}\right]$ post-multiplied by $Z\rceil$ is performed. The resulting column matrix of K + 1 rows is designated by $W\rceil$. The receiver chooses which of the vectors in the $[S]$ matrix was most probably transmitted according to the values of the elements of W. The operation is

$$\left[S_{k+1,n}\right] Z_n\rceil = \left[W_{k+1,1}\right] = W\rceil \quad . \tag{A4-2}$$

Next, one takes a unit orthogonal transformation $\left[A_{n,n}\right]$ such that the vector $\vec{X}_0$ is transformed, for example, to lie along one of the coordinate axes, arbitrarily the first. Then

$$\left[S_{k+1,n}\right]\left[A_{n,n}\right] = \left[S^1_{k+1,n}\right] \quad . \tag{A4-3}$$

In this transformation, it is apparent that

$$\sum_{i=1}^{n} X_{0i} a_{1j} = |X_0| \quad ,$$

while

$$\sum_{i=1}^{n} X_{0i} a_{ji} = 0, \quad j \neq 1 \quad .$$

Further, the sum

$$\sum_{i=1}^{n} a_{ri} a_{si} = \delta^r_s$$

where δ^r_s is the Kronecker delta. The set a_1 of elements will be $a_{1i} = X_{0i}/|X_0|$ while the values of the other elements of A are determined by the orientation of the vector field about the first coordinate axis, which is arbitrary.

Obviously,

$$\left[S_{k+1,n}\right]\left[A_{n,n}\right]\left[A_{n,n}\right]^{-1}\quad Z\Big] = \left[W_{k+1,1}\right] \quad , \tag{A4-4}$$

in which

$$\left[A_{n,n}\right]^{-1} = \left[A_{n,n}\right]_t \quad . \tag{A4-5}$$

But

$$\left[A_{n,n}\right]_t \quad Z\Big] = \begin{bmatrix} |X_0| \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix} + \left[A_{n,n}\right]_t \left[Y\right]_t \quad . \tag{A4-6}$$

Since Y is chosen at random in the n-space in the first place, it might well be chosen after orientation, and thus one may write $Y_n\big]$ for $\left[A_{n,n}\right]_t \left[Y\right]_t$. Then, the final matrix

$$\left[W_{k+1,1}\right] = \left[S'_{k+1,n}\right]\left\{\begin{bmatrix} |X_0| \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix} + \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix}\right\} \quad . \tag{A4-7}$$

A closer inspection of the structure of $\left[S'_{k+1,n}\right]$ reveals the matrix below.

$$\left[S'_{k+1,n}\right] = \begin{bmatrix} |X_0| & 0 \; . \; . \; . \; 0 \\ X'_{11} & X'_{12} \; . \; . \; X'_{1n} \\ . & . \qquad . \\ . & . \qquad . \\ . & . \qquad . \\ X'_{k1} & X'_{k2} \; . \; . \; X'_{kn} \end{bmatrix} \quad . \tag{A4-8}$$

The X'_{ij}'s are also randomly oriented in n-space so that they fit the same probability density distribution after coordinate rotation as before.

In a similar manner, one might choose $\left[A_{n,n}\right]$ to reorient the vector field so that $\vec{Z}$ falls along the first coordinate. In this instance, the combination

$$\left[A_{n,n}\right]_t \quad Z_n\big] = \begin{bmatrix} |Z| \\ 0 \\ . \\ . \\ . \\ 0 \end{bmatrix} \quad , \tag{A4-9}$$

while the combination

$$\left[S_{k+1,n}\right]\left[A_{n,n}\right]=\begin{bmatrix} X''_{01} & X''_{02} & \cdots & X''_{0n} \\ X''_{11} & X''_{02} & \cdots & X''_{1n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ X''_{k1} & X''_{k2} & \cdots & X''_{kn} \end{bmatrix} \cdot \qquad \text{(A4-10)}$$

(Again the numbers X''_{ij} for $i \neq 0$ are from the same density distribution after coordinate rotation as before.) The result finally obtained is given by

$$\left[W_{k+1,1}\right] = |Z| \begin{bmatrix} X''_{01} \\ X''_{11} \\ \cdot \\ \cdot \\ \cdot \\ X''_{k1} \end{bmatrix} \qquad \text{(A4-11)}$$

The result indicated by the matrix $W]$ is the same as that indicated by the column matrix of components of the vectors in the $\vec{Z}$ direction, and thus a decision based on the values of these components is equivalent to one based on the elements of $W]$.

APPENDIX V

THE PROBABILITY DENSITY FUNCTION FOR CORRECT OUTPUTS

The correct correlation output for any of the systems described in which a noise-free version of the reference signal or signals is available to the receiver has, by arbitrary labeling of the results, been designated W_o. This value is given by

$$W_o = |\vec{X}_0|^2 + \vec{X}_0 \cdot \vec{Y} \quad , \tag{A5-1}$$

or

$$W_o = |X_o| \, [|X_o| + Y_1] \quad , \tag{A5-2}$$

where the methods of Appendix IV have been used to reduce the dot product of Eq. (A5-1) to the single algebraic product in Eq. (A5-2).

The probability density distribution function of outputs W_o is designated $p_2(W_o)$, and can be expressed by

$$p_2(W_o) = \int_o^\infty p_o(|X_o|)\, d|X_o| \; \frac{\exp\left[-\dfrac{(W_o - |X_o|^2)^2}{2N\,|X_o|^2}\right]}{\sqrt{2\pi N}\cdot\,|X_o|} \quad . \tag{A5-3}$$

The exponential expression of Eq. (A5-3) is simply the Gaussian distribution function with average $|X_o|^2$ and variance $N|X_o|^2$. Thus, it is the conditional probability function for the value $|X_o|^2 + |X_o| Y_1$ in Eq. (A5-2) subject to given (fixed) $|X_o|$. In Eq. (A5-3), $p_2(W_o)$ is obtained by averaging the conditional distribution over all $|X_o|$.

If the expression for $p_o(|X_o|)$ is inserted in Eq. (A5-2), it becomes

$$p_2(W_o) = \int_o^\infty \frac{2\,|X_o|^{n-1} \exp\left[-\dfrac{|X_o|^2}{2S}\right]}{(2S)^{n/2}\,\Gamma\left(\frac{n}{2}\right)} \cdot \frac{\exp\left[-\dfrac{(W_o - |X_o|^2)^2}{2N\,|X_o|^2}\right]}{\sqrt{2\pi N}\,|X_o|}\, d|X_o| \quad , \tag{A5-3a}$$

which can be simplified somewhat by letting $|X_o|^2 = a$. Then

$$p_2(W_o) = \int_o^\infty \frac{a^{\frac{n+3}{2}} \exp\left[-\dfrac{a}{2S} - \dfrac{(W_o - a)^2}{2Na}\right]}{\sqrt{2\pi N}\,(2S)^{n/2}\,\Gamma\left(\frac{n}{2}\right)} da \quad . \tag{A5-3b}$$

The expression given by Eq. (A5-3b) can be integrated into a complicated expression involving modified Bessel functions of the second kind. However, the characteristic function is more easily obtained and can give a great deal of information about the probability distribution

in an easy manner. Thus

$$\phi_2(t) = \int_{-\infty}^{\infty} \exp[jW_o t]\, p_2(W_o)\, dW_o \quad , \tag{A5-4}$$

$$= \int_{-\infty}^{\infty} p_o(|X_o|)\, d|X_o| \int_{-\infty}^{\infty} \exp[jW_o t]\, \frac{\exp\left[-\frac{(W_o - |X_o|^2)^2}{2N\,|X_o|^2}\right]}{\sqrt{2\pi N}\,|X_o|}\, dW_o \quad . \tag{A5-4a}$$

Let $(W_o - |X_o|^2) = \nu$, then $dW_o = d\nu$; $W_o = \nu + |X_o|^2$.

$$\phi_2(t) = \int_{o}^{\infty} p_o(|X_o|)\, d|X_o|\, \exp[j|X_o|^2 t] \int_{-\infty}^{\infty} \exp[j\nu t]\, \frac{\exp\left[-\frac{\nu 2}{2N\,|X_o|^2}\right]}{\sqrt{2\pi N}\,|X_o|}\, d\nu \quad , \tag{A5-5}$$

$$= \int_{o}^{\infty} p_o(|X_o|)\, d|X_o|\, \exp[j|X_o|^2 t] \exp\left[-\frac{N|X_o|^2 t}{2}\right] \quad , \tag{A5-5a}$$

$$= \int_{o}^{\infty} \frac{2|X_o|^{n-1} \exp\left[-|X_o|^2 \left(\frac{1}{2S} + \frac{Nt^2}{2} - jt\right)\right]}{(2S)^{n/2}\,\Gamma\left(\frac{n}{2}\right)}\, d|X_o| \quad . \tag{A5-5b}$$

The last expression can be easily integrated. If a is again substituted for $|X_o|^2$,

$$\phi_2(t) = \int_{o}^{\infty} \frac{a^{\frac{n}{2}-1} \exp\left[-a\left(\frac{1}{2S} + \frac{Nt^2}{2} - jt\right)\right]}{(2S)^{\frac{n}{2}}\,\Gamma\left(\frac{n}{2}\right)}\, da \quad , \tag{A5-6}$$

so that

$$\phi_2(t) = \frac{1}{(2S)^{n/2}\,\Gamma\left(\frac{n}{2}\right)} \left[\frac{\Gamma\left(\frac{n}{2}\right)}{\left(\frac{1}{2S} + \frac{Nt^2}{2} - jt\right)^{n/2}}\right] \quad , \tag{A5-6a}$$

$$= [1 + SNt^2 - 2jSt]^{-\frac{n}{2}} \quad . \tag{A5-7}$$

From this characteristic function, one may find the moments, and from them obtain a reasonable approximation to $p_2(W_o)$ in normal form. For example,

$$\overline{W_o} = \frac{1}{j}\frac{d}{dt}\phi_2(t)\Big|_{t=o} = \left(\frac{-n}{2j}\right)[1 + SNt^2 - 2jSt]^{-\frac{n+2}{2}}\,(2SNt - 2jS)\Big|_{t=o} \quad , \tag{A5-8}$$

$$= nS \quad . \tag{A5-8a}$$

$$\overline{W_o^2} = -\frac{d}{dt}\left[(1 + SNt^2 - 2jSt)^{-\frac{n+2}{2}} (jnS - nSNt)\right]\Big|_{t=o} \quad , \tag{A5-9}$$

$$= -\left\{(-\frac{n+2}{2})(1 + SNt^2 - 2jSt)^{-\frac{n+4}{2}} (-2n)(SNt - jS)\right.$$

$$\left. + (1 + SNt^2 - 2jS)^{-\frac{n-2}{2}} (-nSN)\right\}\Big|_{t=o} \quad , \tag{A5-9a}$$

$$= nS^2 (n + 2) + nSN \quad . \tag{A5-9b}$$

From the knowledge of $\overline{W_o}$ and $\overline{W_o^2}$, the variance can be computed as

$$\sigma^2_{W_o} = nS^2 (n + 2) + nSN - (nS)^2 \quad , \tag{A5-10}$$

$$\sigma^2_{W_o} = nS (N + 2S) \quad . \tag{A5-10a}$$

Therefore, in as much as $p_2(W_o)$ is the average of a Gaussian distribution with varying average and variance, one may approximate it with the normal density distribution function,

$$p_2(W_o) = \frac{1}{\sqrt{2\pi nS (N + 2S)}} \exp\left[-\frac{(W_o - nS)^2}{2nS (N + 2S)}\right] \quad . \tag{A5-11}$$

This expression is valid chiefly over the bell-shaped part of the curve, and a more precise knowledge of $p_2(W_o)$ is required for an accurate knowledge of the function along its skirts. Returning to Eq. (A5-3b) and rewriting it slightly, one has

$$p_2(W_o) = \int_o^\infty \frac{a^{\frac{n-3}{2}} \exp\left[-\frac{a}{2S} - \frac{W_o^2 - 2W_o a + a^2}{2Na}\right]}{(2S)^{n/2}\,\Gamma(\frac{n}{2})} da \quad , \tag{A5-12}$$

$$= \frac{\exp\left[\frac{W_o}{N}\right]}{(2S)^{n/2}\,\Gamma(\frac{n}{2})} \int_o^\infty a^{\frac{n-3}{2}} \exp\left[-a(\frac{1}{2S} + \frac{1}{2N}) - \frac{W_o^2}{2Na}\right] da \quad . \tag{A5-12a}$$

To evaluate the last given expression, reference to G. N. Watson's "Theory of Bessel Functions" (p. 183, formula 15) reveals

$$\int_o^\infty t^{-\nu-1} \exp\left[-t - \frac{Z^2}{4t}\right] dt = \frac{2K_\nu (Z)}{(\frac{1}{2}Z)^\nu} \quad . \tag{A5-13}$$

To put the integral of Eq. (A5-12a) into the same form, a substitution of t for $a(P/2SN)$ yields

$$\int_0^\infty a^{\frac{n-3}{2}} \exp\left[-a\left(\frac{P}{2SN}\right) - \frac{W_o^2}{4\frac{N^2S}{P}a\frac{P}{2SN}}\right] da$$

$$= \int_0^\infty \left(\frac{2SN}{P}\right)^{\frac{n-3}{2}} t^{\frac{n-1}{2}-1} \exp\left[-t - \frac{PW_o^2}{4N^2St}\left(\frac{2SN}{P}\right)\right] dt \quad , \qquad \text{(A5-14)}$$

$$= \left(\frac{2SN}{P}\right)^{\frac{n-1}{2}} \int_0^\infty t^{-\frac{1-n}{2}-1} \exp\left[-t - \left(\frac{PW_o^2}{N^2S}\right)\frac{1}{4t}\right] dt \quad . \qquad \text{(A5-14a)}$$

Thus, in terms of Eq. (A5-13), $\nu = (1 - n)/2$, and Z is $(W_o/N)\sqrt{P/S}$. Therefore,

$$\int_0^\infty a^{\frac{n-3}{2}} \exp\left[-a\frac{P}{2SN} - \frac{W_o^2}{4Na}\right] da = \left(\frac{2SN}{P}\right)^{\frac{n-1}{2}} \left[\frac{2K_{\frac{1-n}{2}}\left(\frac{W_o}{N}\sqrt{\frac{P}{S}}\right)}{\left(\frac{1}{2}\frac{W_o}{N}\sqrt{\frac{P}{S}}\right)^{\frac{1-n}{2}}}\right] \quad , \qquad \text{(A5-15)}$$

$$= 2\left(\frac{2SN}{P}\right)^{\frac{n-1}{2}} \left(\frac{1}{2N}\sqrt{\frac{P}{S}}\right)^{\frac{n-1}{2}} W_o^{\frac{n-1}{2}} K_{\frac{1-n}{2}}\left(\frac{W_o}{N}\sqrt{\frac{P}{S}}\right) \quad , \qquad \text{(A5-15a)}$$

$$= 2\left(\frac{P}{S}\right)^{\frac{n-1}{4}} W_o^{\frac{n-1}{2}} K_{\frac{n-1}{2}}\left(\frac{W_o}{N}\sqrt{\frac{P}{S}}\right) \quad . \qquad \text{(A5-15b)}$$

Therefore,

$$p_2(W_o) = \frac{2\exp\left[-\frac{W_o}{N}\right] W_o^{\frac{n-1}{2}}}{(2S)^{n/2}\,\Gamma\left(\frac{n}{2}\right)} \left(\frac{P}{S}\right)^{\frac{n-1}{4}} K_{\frac{n-1}{2}}\left(\frac{W_o}{N}\sqrt{\frac{P}{S}}\right) \quad . \qquad \text{(A5-16)}$$

This expression may be used in connection with the limited tables of the Bessel Function if high accuracy is a requirement.

APPENDIX VI

INTEGRATION TO OBTAIN THE PROBABILITY OF ERROR

In Chapter VII, the expression for the probability of error applicable to the system investigated experimentally [see Eq. (7-7)] is

$$P(\text{error}) = \frac{1}{2}\sqrt{\frac{n}{\pi}}\int_0^\infty a^{n-1} \exp\left[-\frac{n}{2}(a^2 - 1)\right] \operatorname{erf}\sqrt{\frac{n\rho}{4a}}\, da \quad . \tag{A6-1}$$

To obtain satisfactory values of P(error), this expression has been integrated numerically by methods designed to give results within about one per cent of the true value of the integral. Such integration is seen to be necessary for small values of n, since the changes in $\operatorname{erf}\sqrt{n\rho/4a}$ are great over the range of values for which $p_o(a)$ is substantially different from zero (see Appendix II).

The integrand of (A6-1) is seen to vanish faster than a simple exponential for large a, and it may be shown that the integrand for very small a is less than $\exp\left[-n/2\,(1-a)^2\right]$. Thus, it is sufficient for the required accuracy to integrate only over those values of a for which the integrand exceeds $\exp[-5]$ times its value at $a = 1$. The value at $a = 1$ is not the maximum value of the integrand, but is of the same order of magnitude.

To determine the limits of a over which integration should be extended, an approximate value for the integral (A6-1) is used. Where $f(n, \rho)$ is used for the probability of error for fixed values of n and ρ, the new form is

$$f(n, \rho) = \sqrt{\frac{2}{\pi}}\int_{-1}^\infty \sqrt{\frac{n}{\pi}} \exp\left[-n\alpha^2\right]\sqrt{\frac{1+\alpha}{n\rho}} \exp\left[-\frac{n\rho}{8(1+\alpha)}\right] d\alpha \quad . \tag{A6-2}$$

The ratio of the integrand to its value at $\alpha = 0$, $(a = 1)$, is given by

$$\text{ratio} = \sqrt{1+\alpha}\, \exp\left[-n\left(\alpha^2 - \frac{\alpha\rho}{8(1+\alpha)}\right)\right] \quad . \tag{A6-3}$$

It is desired to determine the values for which

$$\alpha^2 - \frac{\alpha\rho}{8(1+\alpha)} > \frac{5}{n} \quad . \tag{A6-4}$$

It is obvious that for positive α, $\alpha/(1+\alpha) < 1$, so that if $\alpha^2 = (5/n) + (\rho/8)$ the inequality is satisfied. Furthermore, if α^2 by this equation is less than one, $\alpha/(1+\alpha)$ does not exceed one-half, and the expression for positive α (or upper limit of integration) becomes

$$\alpha = \sqrt{\frac{5}{n} + \frac{\rho}{16}} \quad . \tag{A6-5}$$

The argument leading to Eq. (A6-5) is not valid for α negative. Consider $\beta = -\alpha$, and for $0 < \beta < 1$ it is desired that

$$\beta^2 + \frac{\beta\rho}{8(1-\beta)} > \frac{5}{n} \quad . \tag{A6-6}$$

For small β, a solution to the inequality is

$$\beta = \frac{40}{n\rho} \tag{A6-7}$$

Since the economy of knowing β occurs only when β is small, nothing more need be derived about β.

From Eqs. (A6-5) and (A6-7), one may write

$$P(\text{error}) \doteq \frac{1}{2}\sqrt{\frac{n}{\pi}}\int_{a_1}^{a_2} a^{n-1} \exp\left[-\frac{n}{2}(a^2-1)\right] \operatorname{erf}\sqrt{\frac{n\rho}{4a}}\, da \quad . \tag{A6-8}$$

in which a_1 is the larger of zero of $1 - 40/n\rho$ and a_2 is given by $1 + \sqrt{(5/n)+(\rho/16)}$ unless a_2 exceeds two by this formula, in which case $1 + \sqrt{(5/n)+(\rho/8)}$ should be used.

Below in Table A6-1 are tabulated results of this numerical integration over the limits indicated above. The probability of error is plotted as a function of signal-to-noise-power ratio and time-bandwidth product (or dimension) n in Chapter VII (see Fig. 7.2).

TABLE AIV-1

PROBABILITY OF ERROR FOR THE BINARY CHOICE SYSTEM OBTAINED BY NUMERICAL INTEGRATION

n ↓ \ ρ →	0.001	0.01	0.1	1.0	10.0
10	0.480*	0.433	0.307	0.0548	1.94×10^{-6}
10	0.437*	0.308	0.0568	3.09×10^{-7}	1.30×10^{-56}*
10	0.309*	0.0566	2.59×10^{-7}	1.30×10^{-56}*	0
10	0.0569*	2.87×10^{-7}	1.30×10^{-56}*	0	0
10	2.88×10^{-7}*	1.30×10^{-56}*	0	0	0

* = entries taken from the error function table of $1/2 \operatorname{erf}\sqrt{n\rho/4}$ in which n is considered sufficiently large that the change in value of the error function over the range of contribution of a is small.

The 0's are inserted in those positions of the table where the probability of error is insignificant in the lifetime of an equipment.

APPENDIX VII

SCHEMATICS OF THE EXPERIMENTAL NOMAC SYSTEM

The block diagram below (Fig. A7.1) is a guide to the complete schematic diagrams on the following pages.

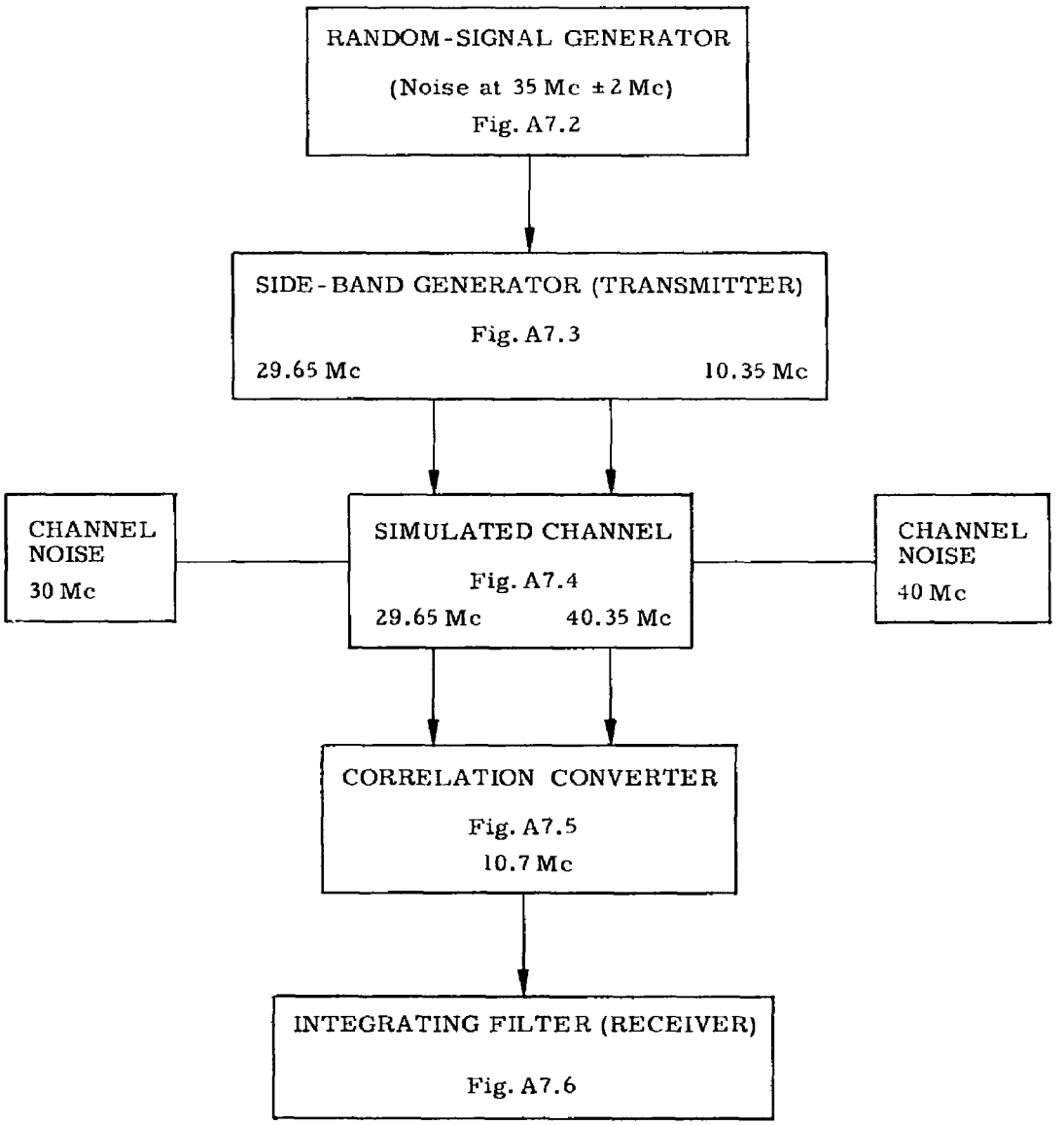

Fig. A7.1. Block diagram of experimental system.

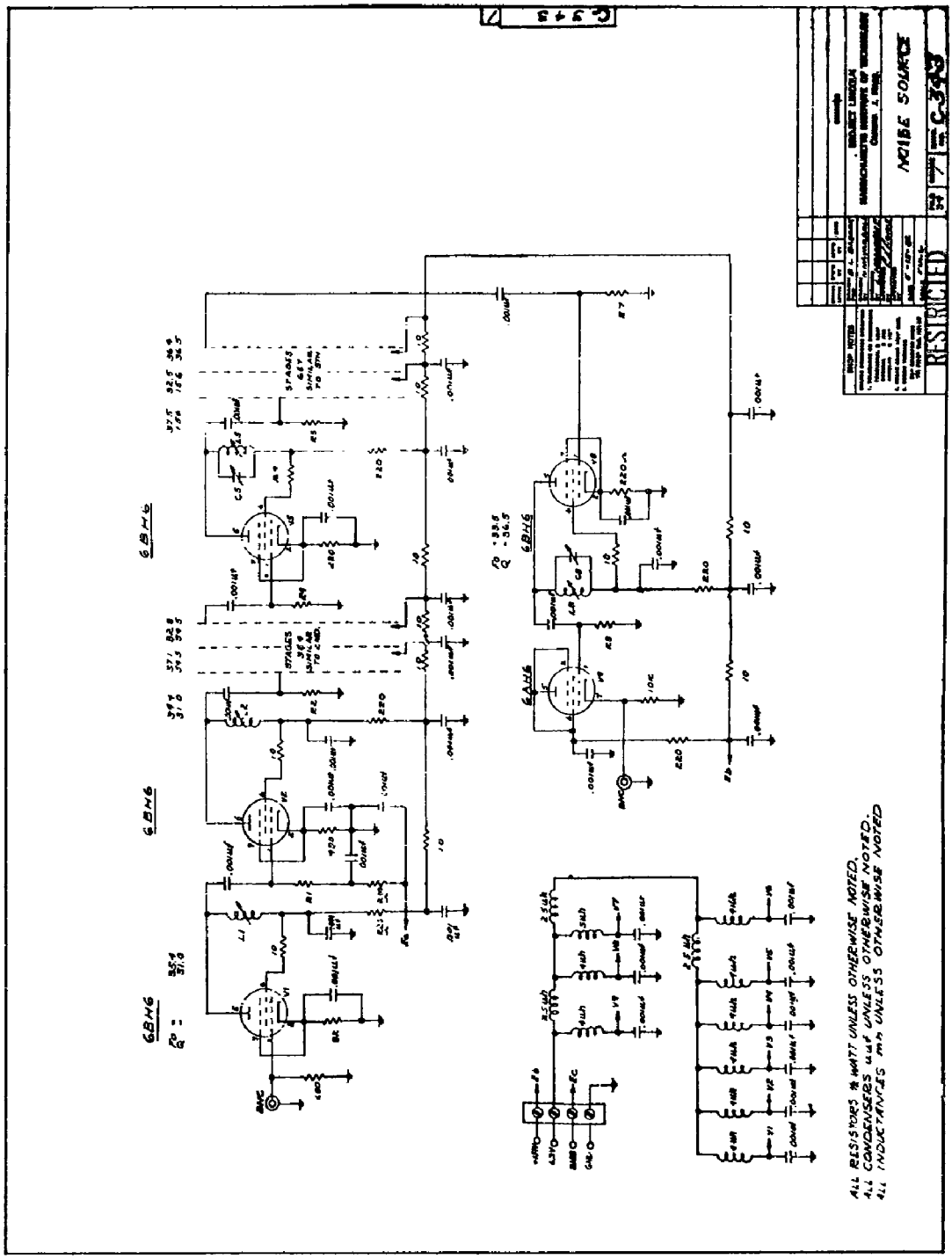

Fig. A7.2. Schematic diagram of random-signal generator.

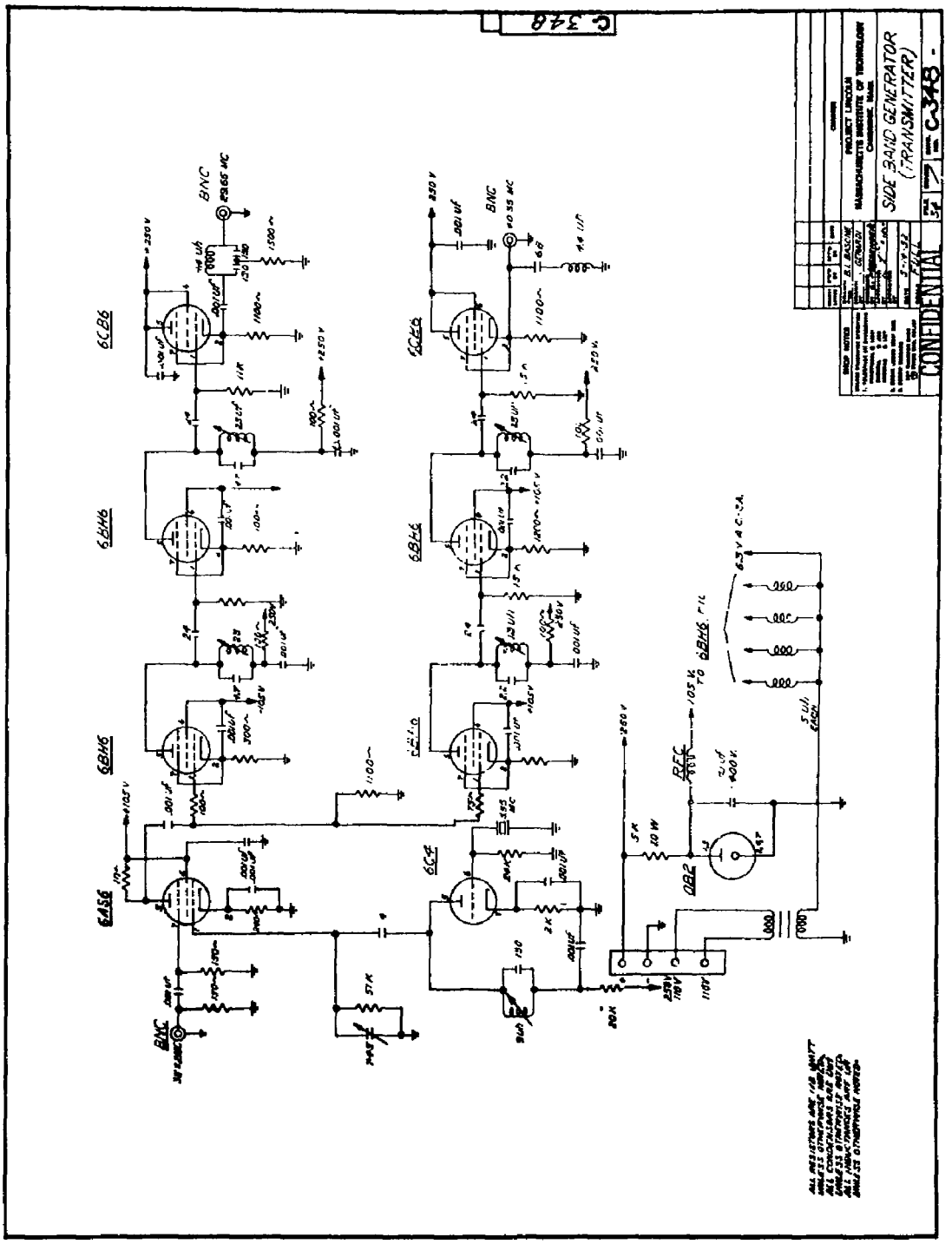

Fig. A7.3. Schematic diagram of side-band generator (transmitter).

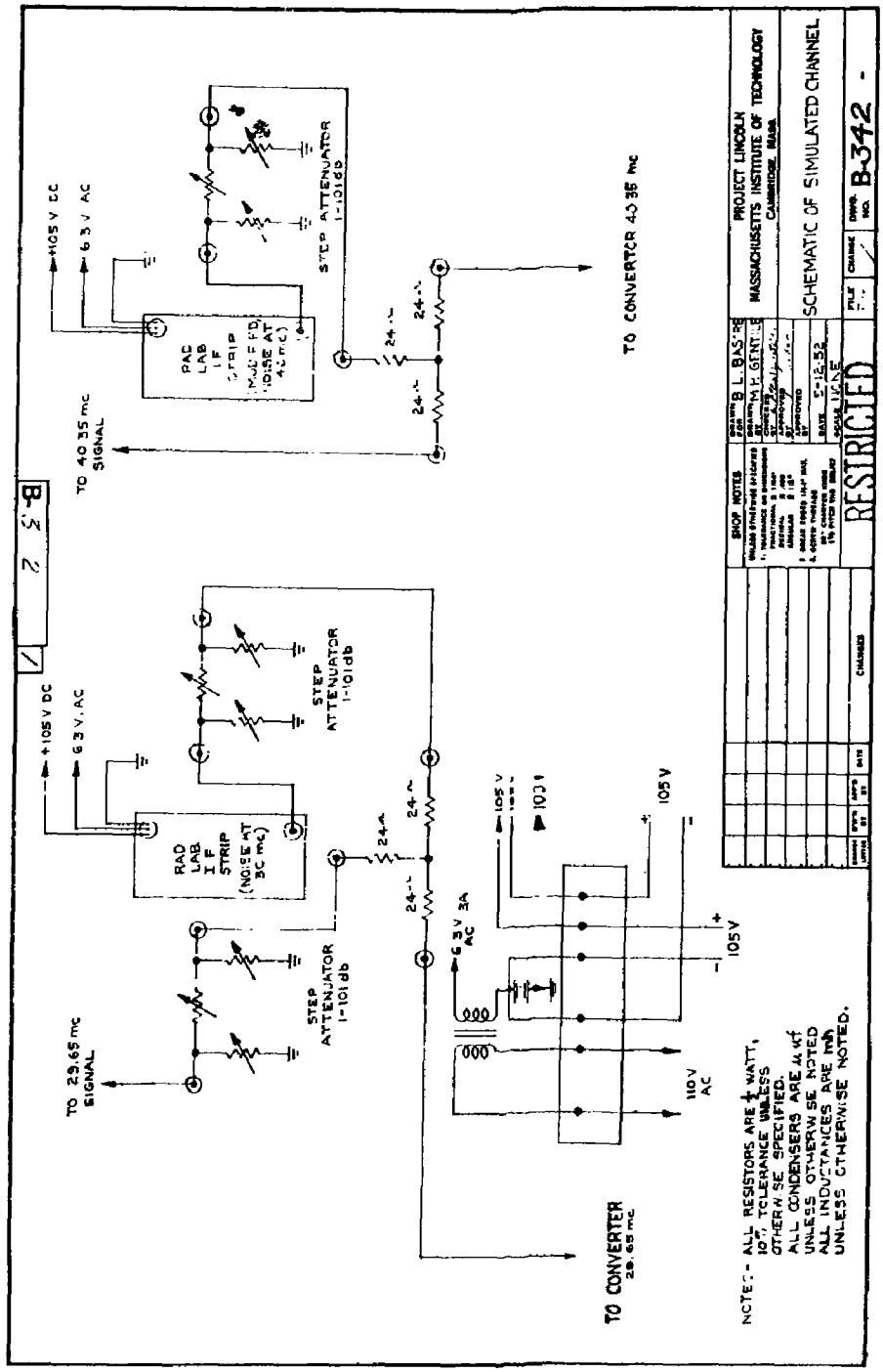

Fig. A7.4. Schematic diagram of simulated channel.

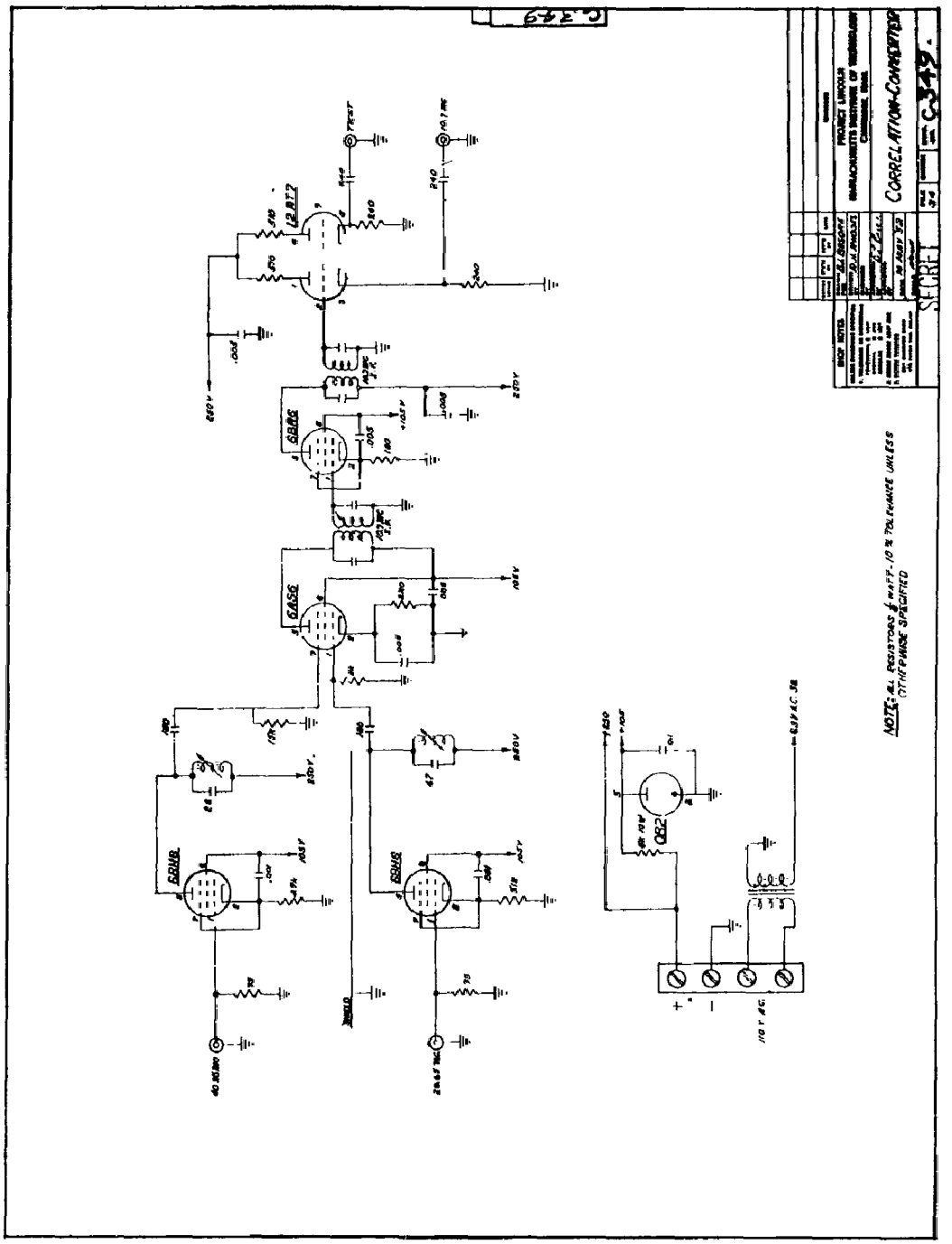

Fig. A7.5. Schematic diagram of correlation converter.

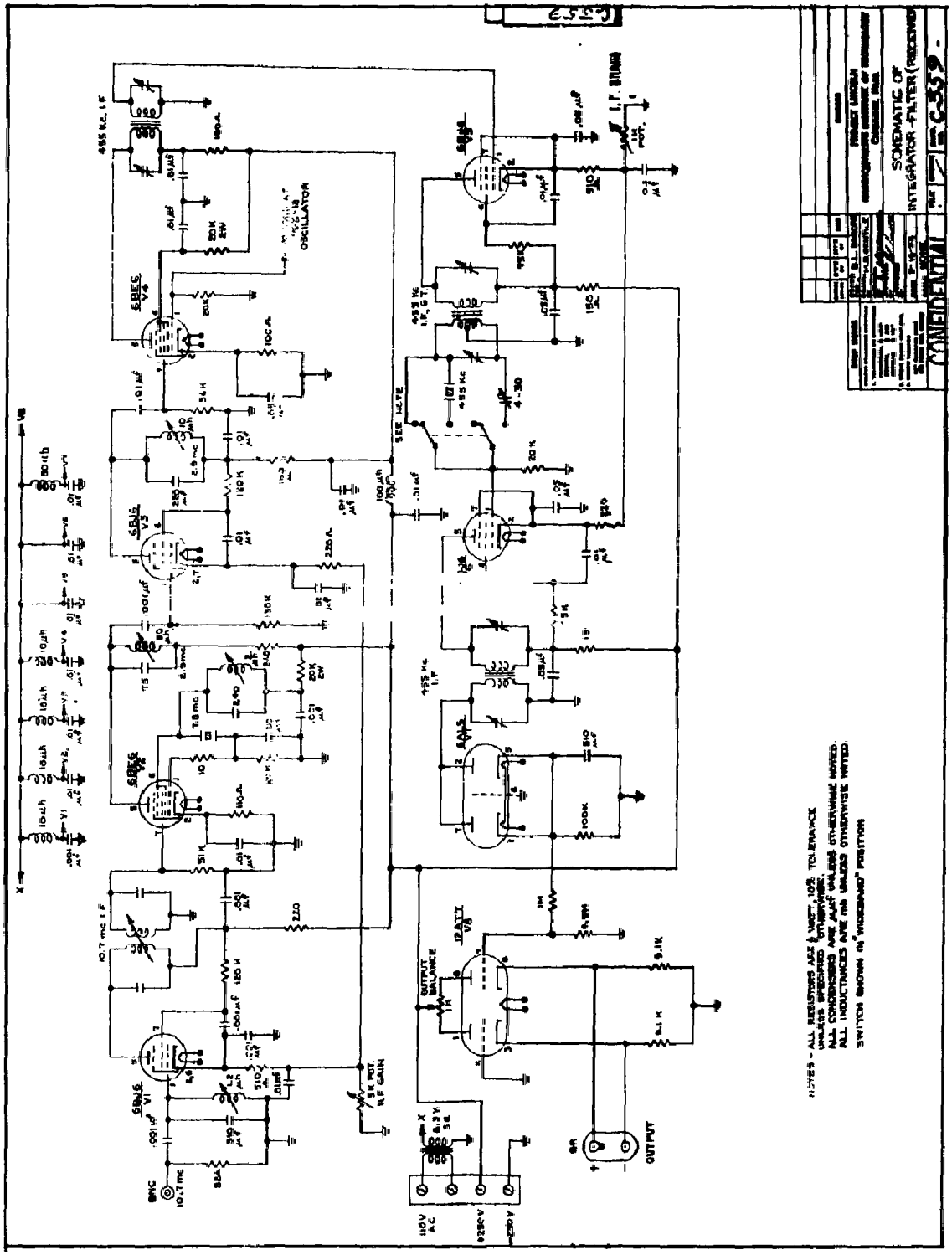

Fig. A7.6. Schematic diagram of integrating filter (receiver).

APPENDIX VIII

THE DISTRIBUTION OF SUMS OF PRODUCTS

Although the probability density distribution discussed in this Appendix does not appear in the main body of the paper, it was extensively investigated during the preliminary research work. Since it is inherently related to the correlation process, a review of the results is given here.

Let U be defined as the sum of products XY where X and Y are independent normal variates, the variance of the former being S and that of the latter being N. Thus,

$$U = \sum_{i=1}^{n} U_i = \sum_{i=1}^{n} X_i Y_i \quad . \tag{A8-1}$$

The joint distribution of X_i and Y_i is written from a knowledge of their individual distributions.

$$p(X, Y) = \frac{\exp\left[-\frac{X^2}{2S} - \frac{Y^2}{2N}\right]}{2\pi\sqrt{SN}} \quad . \tag{A8-2}$$

The probability density distribution for $U_i = X_i Y_i$ is given by

$$p'(|U_i|) = \int_0^\infty \frac{2}{\pi\sqrt{SN}\,a} \exp\left[-\frac{a^2}{2S} - \frac{U_i^2}{2Na^2}\right] da \quad , \tag{A8-3}$$

where the integration is carried out only in one quadrant of the joint distribution, which is symmetrical about both X and Y axes. As given in (A8-3) p' is an even function of U_i on the left, and, since positive or negative products occur with equal likelihood, it is apparent that

$$p(U_i) = \frac{1}{\pi\sqrt{SN}} \int_0^\infty \frac{1}{a} \exp\left[-\frac{a^2}{2S} - \frac{U_i^2}{2Na^2}\right] da \quad . \tag{A8-4}$$

To obtain the desired density distribution function, the characteristic function is found as an intermediate step.

$$\phi'(t) = \int_{-\infty}^{\infty} \exp[itU_i] \int_0^\infty \frac{1}{\pi a\sqrt{SN}} \exp\left[-\frac{a^2}{2S} - \frac{U_i^2}{2Na^2}\right] da\, dU_i \quad , \tag{A8-4}$$

$$= \int_0^\infty da \frac{\exp\left[-\frac{a^2}{2S}\right]}{\pi a\sqrt{SN}} \int_{-\infty}^{\infty} \exp\left[-\frac{U_i^2}{2Na^2} + itU_i\right] dU_i \quad , \tag{A8-4a}$$

$$= \int_0^\infty \frac{2}{\sqrt{\pi S}} \exp\left[-\frac{a^2}{2}\left(\frac{1}{S} + N^2 t\right)\right] da \quad , \tag{A8-4b}$$

$$= [1 + SNt^2]^{-\frac{1}{2}} \quad . \tag{A8-5}$$

The characteristic function of the desired probability density distribution function p(U) is then given by

$$\phi(t) = [1 + SNt^2]^{-\frac{n}{2}} \quad , \tag{A8-6}$$

and from it,

$$p(U) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp[-itU]\,[1 + SNt^2]^{-\frac{n}{2}}\,dt \quad . \tag{A8-7}$$

From the characteristic function, one obtains a complete description of the distributions of U's. For example, from the expansion of $\phi(t)$, one obtains $\overline{U^2} = nSN$ and thus, $\sigma_U^2 = nSN$, since $\overline{U}$ is zero.

Foster and Campbell[2] list the integral (A8-7) in their comprehensive table of transforms. In particular, transfer pair No. 569 links $1/[\beta^2 + \rho^2]^{\alpha}$ and

$$\frac{|g|^{\alpha-\frac{1}{2}}\, K\alpha - \frac{1}{2}(\beta|g|)}{\sqrt{\pi}\,\Gamma(\alpha)\,(2\beta)^{\alpha-\frac{1}{2}}} \quad .$$

When this is applied to Eq. (A8-7), the result is

$$p(U) = \frac{\left(\frac{U}{2\sqrt{SN}}\right)^{\frac{n-1}{2}} K_{\frac{n-1}{2}}\left(\frac{U}{\sqrt{SN}}\right)}{\sqrt{\pi SN}\,\Gamma\left(\frac{n}{2}\right)} \quad . \tag{A8-8}$$

The $K_\nu(z)$ here is the modified Bessel Function of the second kind for imaginary arguments of order ν. Many sources (e.g., Hildebrand[13]) give the behavior of

$$K_\nu(z) \xrightarrow[z\to 0]{} 2^{\nu-1}\,\Gamma(\nu)\,z^{-\nu} \quad .$$

When this is substituted in Eq. (A8-8), the value of p(U = 0) is obtained, and is found to be $1/\sqrt{2NSn\pi}$.

An interesting way of examining p(U) can be developed as follows. Let

$$\phi(t) = \phi_a(t)\,\phi_b(t) \quad .$$

Thus,

$$\phi_a(t) = \frac{1}{[1 + i\sqrt{SN}\,t]^{n/2}} \quad , \tag{A8-9}$$

and

$$\phi_b(t) \quad \frac{1}{[1 - i\sqrt{SN}\,t]^{n/2}} \quad . \tag{A8-9a}$$

Since it is known that p(U) is the transform of $\phi(t)$, it must be the convolution of two distribution functions $p_a(V)$ and $p_b(W)$, namely,

$$p(U) = \int_{-\infty}^{\infty} p_a(V)\, p_b(U - V)\, dV \quad , \tag{A8-10}$$

in which

$$p_a(V) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp[-itV]\, \phi_a(t)\, dt \quad , \tag{A8-10a}$$

and

$$p_b(W) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp[-itW]\, \phi_b(t)\, dt \quad . \tag{A8-10b}$$

But integrals of the type (A8-10a) and (-10b) were studied in Appendix II. In particular, it was found that $p_b(W)$ is of the form of $p_o(W)$ in which (2S) has been replaced by $\sqrt{SN}$, and similarly, $p_a(V) = p_o(-V)$ with the same substitution. Thus, W exists only as a positive number and V is a negative number. The variable desired, U, is a sum of a positive and a negative number taken from the same distribution of magnitudes (known as the "chi-squared" distribution in statistics). Thus, the form of Eq. (A8-10) may be changed to read

$$p(U) = \int_{-\infty}^{\infty} p_o(W)\, p_o(U + W)\, dW \quad . \tag{A8-11}$$

This is recognized as the identical form of the autocorrelation function of U for the probability density distribution p_o!

When n is even (n - 1/2 is half an odd integer) the Bessel Function reduces to a polynomial. This fact can be obtained as a result of the autocorrelation process indicated in (A8-11) – namely, if $p_o(W) = \alpha W^{p-1} \exp[-\beta W]$, α and β being written for more complicated constants,

$$p(U) = \int_{o}^{\infty} \alpha^2 W^{p-1} (U + W)^{p-1} \exp[-2\beta W - \beta U]\, dW, \; U \geq 0 \quad , \tag{A8-12}$$

$$= \alpha^2 \exp[-\beta U]\int_{o}^{\infty} W^{p-1} \sum_{k=o}^{p-1} {}_{p-1}C_k\, W^{p-k-1}\, U^k \exp[-2\beta W]\, dW \quad . \tag{A8-12a}$$

The order of integration and summation may be interchanged, and the integration resulting is easy to perform. Then,

$$p(U) = \frac{\alpha^2 \exp[-\beta U](p-1)!}{(2\beta)^{2p-1}} \sum_{k=o}^{p-1} \frac{(2p-K-2)!}{(p-K-1)!K!} (2\beta)^k U^k, \; U \geq 0 \quad . \tag{A8-13}$$

When α and β are replaced by their values in terms of S and N,

$$p(U) = \frac{\exp -\frac{|U|}{\sqrt{SN}}}{n(\frac{n}{2})!\; 2^n \sqrt{SN}} \sum_{k=o}^{\frac{n}{2}-1} \frac{(n-2-K)!}{(\frac{n-2}{2}-K)!K!} \left(\frac{2|U|}{\sqrt{SN}}\right)^k \quad , \tag{A8-14}$$

valid for all U when n is even. The absolute-value symbols are included because (A8-12) is derived only for positive U. The knowledge of the property of the autocorrelation function, i.e.,

$p(U) = p(-U)$, is used to extend the result to negative values of U.

The first check on (A8-14) occurs for $n = 2$, in which case

$$p(U) = \frac{1}{2\sqrt{SN}} \exp\left[-\frac{|U|}{\sqrt{SN}}\right] .$$

This may be readily verified by integration of (A8-7) for this value of n.

As a final inspection of the properties of p(U), it is interesting to compare certain properties of the density distribution function with those of the function of Appendix II, $p_o(X)$. In particular, if $\sqrt{2S}$ is replaced by a in p_o and $\sqrt{NS}$ is replaced by a in p(U), it develops that the maximum value of each distribution function is $1/a\sqrt{2\pi n}$, while the variance associated with each is na^2. They differ in their other moments, however, in as much as $p_o(X)$ is defined only for magnitudes, and is centered about an average $na/\sqrt{2}$, while the average value of U is zero.

AD [redacted]

4641

SECRET